1831 1952

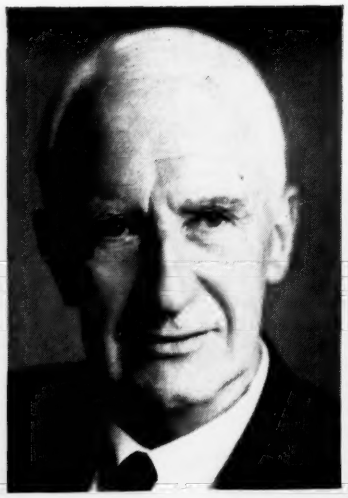
THE ADVANCEMENT OF SCIENCE

VOL. IX, No. 34 SEPTEMBER, 1952

CONTENTS

Addresses delivered at the Annual Meeting of the British Association for the Advancement of Science in Belfast, September 3 to 10, 1952.

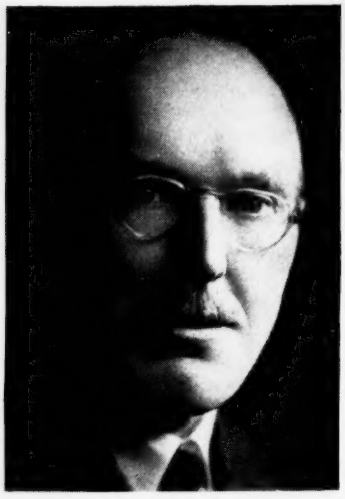
	PAGE
PORTRAITS OF THE PRESIDENT OF THE ASSOCIATION AND OF PRESIDENTS OF THE SECTIONS	90
OFFICERS AND COUNCIL	92
THE ETHICAL DILEMMA OF SCIENCE, *by* PROF. A. V. HILL, C.H., O.B.E., F.R.S.	93
SOME CONTRIBUTIONS TO EXPERIMENTAL PHYSICS FROM BRITAIN, *by* PROF. A. M. TYNDALL, C.B.E., F.R.S.	103
THE ADVANCING FRONT OF CHEMISTRY, *by* PROF. W. WARDLAW, C.B.E.	113
SOME ASPECTS OF GEOLOGICAL RESEARCH AND THEIR PRACTICAL APPLICATION, *by* DR. J. E. RICHEY, M.C., F.R.S.	122
SOME PROBLEMS OF PARTHENOGENESIS, *by* PROF. A. D. PEACOCK	134
SOME ASPECTS OF SETTLEMENT IN THE OVERSEAS DOMINIONS, *by* PROF. R. OGILVIE BUCHANAN	149
FULL EMPLOYMENT IN RETROSPECT, *by* PROF. D. T. JACK, C.B.E.	159
PROGRESS IN AERONAUTICAL SCIENCE AND ENGINEERING, *by* SIR BEN LOCKSPEISER, K.C.B., F.R.S.	171
ARCHÆOLOGY, FOLKLORE AND SACRED TRADITION, *by* PROF. E. O. JAMES	184
THE NERVOUS CONTROL OF THE MOVEMENTS OF THE ALIMENTARY CANAL, *by* PROF. R. C. GARRY	197
THE ASSESSMENT OF PERSONALITY, *by* PROF. P. E. VERNON	207
THE ANATOMY OF METABOLISM IN HIGHER PLANTS: TOTAL METABOLISM, METABOLIC PROCESSES AND PARTIAL REACTIONS, *by* PROF. MEIRION THOMAS, F.R.S.	219
EDUCATION OF AFRICANS IN EAST AND CENTRAL AFRICA, *by* A. L. BINNS, C.B.E., M.C.	236
BRITAIN'S FOOD SUPPLIES, *by* DR. N. C. WRIGHT	245



Prof. A. V. Hill
1.

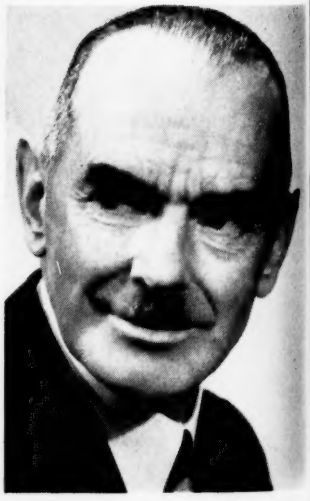
Prof. A. M. Tyndall
2.

Prof. W. Wardlaw
3.



Dr. J. E. Richey
4.

Prof. A. D. Peacock
5.



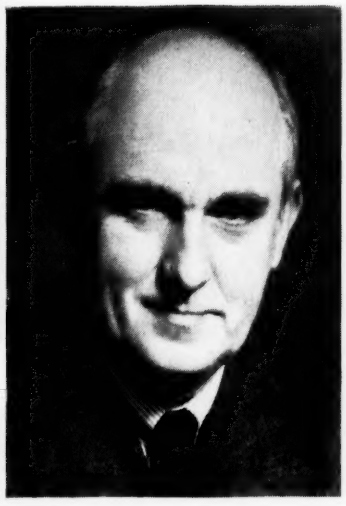
Prof. R. O. Buchanan
6.

Prof. D. T. Jack
7.

Portraits 1, 2, 8 *and* 12 *by Walter Stoneman, London.*

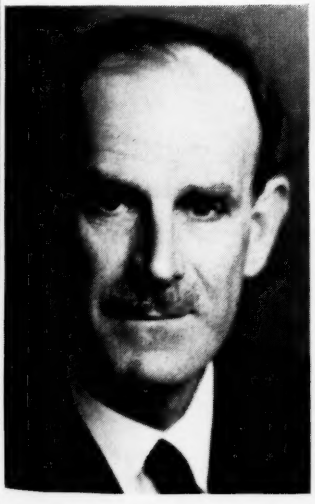
Sir Ben Lockspeiser
8.

Rev. Prof. E. O. James
9.



Prof. R. C. Garry
10.

Prof. P. E. Vernon
11.



Prof. Meirion Thomas
12.

Mr. A. L. Binns
13.

Dr. N. C. Wright
14.

Portraits 9 and 7 by Elliott & Fry Ltd. London.

OFFICERS AND COUNCIL OF THE BRITISH ASSOCIATION, 1952-3

PATRON.

HER MAJESTY THE QUEEN.

PRESIDENT, 1952.

Prof. A. V. Hill, C.H., O.B.E., F.R.S.

PRESIDENT-ELECT, 1953.

Sir Edward Appleton, G.B.E., K.C.B., F.R.S.

GENERAL TREASURER.

M. G. Bennett.

GENERAL SECRETARIES.

Sir Richard Southwell, F.R.S. | Dr. George Taylor.

ORDINARY MEMBERS OF THE COUNCIL.

Prof. F. Balfour-Browne.
R. J. Bartlett.
Prof. L. F. Bates, F.R.S.
M. C. Burkitt.
Wing-Comm. T. R. Cave-Browne-Cave, C.B.E.
Prof. H. G. Champion, C.I.E.
Prof. Winifred Cullis, C.B.E.
Sir Alfred Egerton, F.R.S.
Prof. H. P. Gilding.
Dr. Ezer Griffiths, O.B.E., F.R.S.
Dr. D. B. Harden.
Prof. K. S. Isles.
A. Gray Jones.
Sir Harold Spencer Jones, F.R.S.
Prof. H. D. Kay, C.B.E., F.R.S.
Sir Ben Lockspeiser, K.C.B., F.R.S.
J. Wickham Murray, O.B.E.
Prof. F. W. Paish.
Prof. W. J. Pugh, O.B.E., F.R.S.
Dr. J. Ramsbottom, O.B.E.
Lord Rennell of Rodd, K.B.E., C.B.
Prof. L. Dudley Stamp, C.B.E.
Dr. W. E. Swinton.
Dr. H. Hamshaw Thomas, M.B.E., F.R.S.
Dr. C. Tierney.
Prof. W. Wardlaw, C.B.E.
W. H. Wilcockson.
Dr. H. E. Wimperis, C.B., C.B.E.
Prof. A. W. Wolters.
Dr. N. C. Wright.

EX-OFFICIO MEMBERS OF THE COUNCIL.

Past-Presidents of the Association, the President for the year, the President and Vice-Presidents for the ensuing Annual Meeting, past and present General Treasurers and General Secretaries, and the Local Treasurers and Local Secretaries for the Annual Meetings immediately past and ensuing.

PAST PRESIDENTS OF THE ASSOCIATION.

H.R.H. The Duke of Windsor, K.G., F.R.S. (1926).
Prof. Sir Arthur Keith, F.R.S. (1927).
Sir Henry Dale, O.M., G.B.E., F.R.S. (1947).
Sir Henry Tizard, G.C.B., F.R.S. (1948).
Sir John Russell, O.B.E., F.R.S. (1949)
Sir Harold Hartley, K.C.V.O., C.B.E., M.C., F.R.S. (1950).
H.R.H. The Duke of Edinburgh, K.G., K.T., F.R.S. (1951).

PAST GENERAL OFFICERS OF THE ASSOCIATION.

Prof. P. G. H. Boswell, O.B.E., F.R.S.
Prof. Sir David Brunt, Sec. R.S.
Sir Harold Hartley, K.C.V.O., C.B.E., M.C., F.R.S.
Dr. Edward Hindle, F.R.S.
Sir John Lennard-Jones, K.B.E., F.R.S.
Prof. Sir John L. Myres, O.B.E.
Sir Frank Smith, G.C.B., G.B.E., F.R.S.
Prof. F. J. M. Stratton, O.B.E., F.R.S.
Sir Arthur Trueman, K.B.E., F.R.S.

HON. AUDITORS.

Sir John Simonsen, F.R.S. | Dr. H. J. T. Ellingham.

SECRETARY.

D. N. Lowe, O.B.E.

THE ETHICAL DILEMMA OF SCIENCE

Address by

PROF. A. V. HILL, C.H., O.B.E., F.R.S.

PRESIDENT OF THE ASSOCIATION

EXACTLY a hundred years ago the British Association was meeting for the first time in Belfast ; we are happy indeed to be gathered here today, in this hospitable city, to test your hospitality again and celebrate with you a century of progress of British science. The part which your kinsfolk have played in this makes a long and honourable story. Often their work was done in other parts of our Islands, or far away overseas : that belongs to your ungrudging tradition of service and adventure, and it does not stop you from producing, in every generation, plenty more of their kind. In 1660 the little town of Killyleagh, some twenty miles from here, was the birthplace of Hans Sloane, botanist, collector, physician, and president of the Royal Society, on whose bequest to the nation the British Museum in London was founded : its two-hundredth anniversary will be celebrated next year. In the following century Joseph Black, the son of a native of Belfast and educated here, was a pioneer in chemistry and the theory of heat. In 1824, in this city, William Thomson was born, Lord Kelvin, the famous physicist and inventor, his father a native of Co. Down being then a teacher of mathematics in the Royal Academical Institution here. Joseph Larmor, born in Co. Antrim in 1857, was taught at that school, and when he became Senior Wrangler at Cambridge in 1880, repeating the success of a schoolmate the year before, a torchlight procession was held in Belfast—to the bewilderment (as Eddington records) of the shy young mathematician. This city was the birthplace in 1866 of E. W. MacBride, the zoologist, also at the same school : those of us who remember him affectionately could never have imagined him coming from anywhere else. A dear friend of many of us, Joseph Barcroft the physiologist, was born at Newry in 1872, and learnt what he knew of the things that really matter (as he truly said) sailing in Carlingford Lough. And among the living, Frederick Donnan, a native of Holywood, Co. Down, and a graduate of Queen's, whose imagination has fertilised both chemistry and biology, tells me that he took part in the meeting here in 1902. Thus it goes on. British science owes much to the fertile and imaginative minds, the vigorous temperaments and the warm hearts of your kinsmen ; and in celebrating the centenary of our first visit to Belfast we and you can remember them with pride.

We can recall too that applied science is a partner in the great industries for which Belfast and Northern Ireland are famous, the city for shipbuilding, engineering and textiles, the country around for agriculture. In the Report of the British Association for 1852 there is a description of the vortex water wheel, an early form of turbine, in working use at a near-by mill : James Thomson, elder brother of Lord Kelvin and for many years professor of civil engineering at Queen's College, had patented a vortex water wheel in 1850. In the same Report is a long account, by the Professor of Agriculture at Queen's College, of the composition and economy of the flax plant : while thirty pages were devoted to the fattening of animals—English animals, it is true, but their cousins here have been apt pupils. In such practical arts, based alike on scientific knowledge and traditional skill, the contribution of Northern Ireland has been as distinguished as in the advancement of science itself ; and we may confidently expect that the present meeting of the British Association will serve not only to celebrate the achievements of the last century, but to foster the endeavours of the next.

The President of 1852 was Edward Sabine, astronomer, explorer and geodesist, a colonel then in the Royal Artillery. Sabine at that time was treasurer of the Royal Society, at various others he was physical secretary, foreign secretary and president : indeed, for completeness, he should have been biological secretary too, for his name occurs in an index of British and Irish Botanists, and in the British Museum are plants collected by him in Arctic expeditions. Sabine was an Irishman, and if your president today is English, with an ancestor from Northern Ireland two hundred years ago, that only emphasises, after a century, the unity of science. Indeed, there are many Irishmen with us today from the other side of the border : they are doubly welcome, for science admits no frontiers, and customs duty is not levied on ideas or friendship. Sabine's address referred particularly to the subject of his own chief interest for many years, the periodic variations of terrestrial magnetism : indeed, he announced in his address the discovery of the connexion between sunspots and magnetic disturbances in the earth. He finished his address by referring to ' allusions . . . made by influential men . . . to a direct representation of Science in Parliament.' The benefit, he said, which the Legislature might derive from such a change was a question rather for statesmen than for scientists ; but as regards Science itself he expressed his strong conviction that the possible gain would be far outweighed by inevitable loss, and that scientific men could not too highly value the advantage they possessed in the undisturbed enjoyment of their own pursuits untroubled by the excitements and distractions of political life. The practical importance of Science today, and its impact on public affairs, have greatly reduced that undisturbed enjoyment ; and though all would agree that the direct representation of Science as such in the House of Commons is impracticable, none would doubt the advantage to Parliament and the nation if more of its members had some personal acquaintance with science. It might indeed be well, in a reformed Second Chamber, to provide the same representation to Science as at present is afforded to the Church and the Law.

Two famous young Irishmen were present at the meeting here in 1852. William Thomson, Lord Kelvin, elected to the Chair of Natural Philosophy at Glasgow six years earlier at the unusual age of 22, was president of the Section of Physics : the other, George Gabriel Stokes, then 33, delivered a public lecture on fluorescence due to ultraviolet light passing through a solution of quinine. Both were among the foremost physicists of the following half century, and Kelvin succeeded Stokes as president of the Royal Society.

The next meeting in Belfast was in 1874, when another Irishman, John Tyndall, presided. His notable address will be referred to later. Huxley delivered an evening Discourse with the provocative title, ' The hypothesis that animals are automata ' ; whether for that or for some other cause it is recorded that he was threatened with assault by a nonconformist minister. On the lighter side, a description was given of an old lady of 111 who chatted away continually in a clear voice, was in possession of all her faculties, took snuff, and had white hair and a skin as solf as velvet : which is encouraging to those who take snuff and have white hair.

The last meeting at Belfast was 50 years ago, in 1902, under the presidency of James Dewar, a Scotsman, famous for his work on the liquefaction of gases. Lord Kelvin, loyal to Belfast, read a curious paper in the Physics Section entitled ' Animal Thermostat.' The great physicist, arguing on thermodynamic principles, speculated whether the breath of an animal, kept a considerable time in a hot bath above the natural temperature of its body, might be found to contain no carbon dioxide at all ; possibly even a surplus of oxygen, pointing to an ' unburning ' of matter in the body. I hope the physiologists were present at the discussion ; one of them, Joseph Barcroft, would certainly have offered to make the experiment on himself. Henry Armstrong, provocative as usual and at considerable length, gave a presidential address to the Section of Educational Science. Today, fifty years after, we are delighted to have with us a few of those who were here in 1902 : we bid them a warm welcome to their jubilee.

The Second War delayed our coming

again ; after which, Belfast, licking her honourable wounds, needed time to get ready. But here we are assembled, after 50, 78 and 100 years. There is every promise of a notable meeting, and the certainly of a most friendly and hospitable reception by the city and citizens, and the Queen's University, of Belfast. There will be other occasions for thanks, but at our Inaugural Meeting it is fitting that the gratitude of the Association should be expressed at once to all who have worked so hard to make our gathering happy and successful.

It is a very special pleasure to be called to the Presidency of the Association in succession to H.R.H. the Duke of Edinburgh, who won the confidence and esteem of scientific people at our meeting last summer, confirmed by many contacts with them since in Canada and the United States and recently in Britain. He presided at the December meeting of the Council, and we told him then how glad we were that these contacts were soon to be extended to Ceylon, Australia and New Zealand. Those hopes, alas, have had to be deferred ; and on behalf of the Association I wrote in February to tell him of our deep sympathy in the bereavement which H.M. The Queen and he had suffered. We mourned the loss, not only of a Royal Patron but of one whom all regarded with personal gratitude and affection. I added, and I know you will warmly confirm, that our sorrow and sympathy were accompanied, since we had seen him at work with us, by the certain expectation that the same confident affection will support The Queen and him in the years ahead.

The Duke of Edinburgh concluded his Presidential Address last year with the words,

> ' It is clearly our duty as citizens to see that science is used for the benefit of mankind. For, of what use is science if man does not survive ? '

Here was a challenge to his successor : to discuss how far science has already contributed to human betterment, how far it has provided fresh problems, dangers and difficulties ; and to suggest ways in which all who are concerned with science can help, as citizens, to make sure that its results in fact are beneficial.

As citizens : for scientists as such have no title to superior wisdom or virtue, and outside their special knowledge they are just as likely as others to be misled. The fundamental principle of scientific work is unbending integrity of thought, following the evidence of fact wherever it may lead, within the limits of experimental error and honest mistake. On this there can be no compromise. And since science is a universal interest of mankind, recognising no barriers of race, class, religion, or opinion (provided that is honest), a necessary condition of its advance and application is one of friendliness, frankness and equality. Goodwill and integrity, therefore, are indispensable alike to scientific progress itself and its successful employment for the benefit of mankind. Those who look to scientists as magicians, able to conjure a universal formula out of a hat, may be disappointed to find only so ancient a doctrine : and admittedly there *is* far more to science than integrity and goodwill. But these are the qualities chiefly required to utilise the opportunities, to resolve the problems and difficulties, which science has provided for present-day society.

The common phrase, ' this scientific age,' is all too apt to imply, with little justification, that the majority of people, at least in highly developed countries, now think and act scientifically ; and, with no justification at all, to suggest that science can replace the older motives of human conduct. It is true that the external circumstances of life have been vastly altered by the applications of scientific discovery and invention, though as yet for only a minority of mankind. The future alone can decide whether natural resources and human ingenuity will prove sufficient, given statesmanship and goodwill, for the same transformation gradually to affect the whole of human society. If not, are stable conditions ultimately possible ? Or will there be perpetual conflict between the ' haves ' and the ' have-nots '? It is true also that the methods, ideas and results of scientific inquiry have penetrated widely, if not deeply, into popular thinking and belief : the jargon at least of science is widespread, and magic and superstition are gradually losing, if not their currency at least their respectability. Yet such

changes may have little real influence on the basic pattern of human behaviour, and if witches are no longer hunted down and killed, political and racial intolerance can lead to even wilder and crueller excesses. In clearing away old idolatries there is always a danger of allowing new ones to creep in : the unclean spirit went out when the house was swept and garnished, but only to return with seven others more wicked than himself. The improvement of man's estate by the application of scientific knowledge is one of the loftiest of adventures : but a belief that it can be achieved by scientific methods alone, without a moral basis to society, is a perilous illusion. If the methods of human experiment and racial improvement adopted by the Nazis could be regarded purely as applied biology, there might be much to say for them. But most of us believe that by abandoning a faith (which has nothing directly to do with science) in the sanctity of the human individual and of moral law, they were heading straight for disaster. Yet we shall see later the dilemma in which such scruples put us, in respect of the gravest of all world problems.

The conflict between new knowledge and traditional belief is no novelty. When Eve saw that the tree of knowledge was good for food and that it was pleasant to the eyes, and to be desired to make one wise, she took of the fruit thereof and did eat and gave also to Adam and he did eat. Which led, as you have read, to their exclusion from the garden and the warning ' in sorrow shalt thou eat of it all the days of thy life.' Again and again the attempt has been made to forbid the fruit of scientific knowledge. In 1874 in this city John Tyndall delivered a presidential address to the British Association which provoked a hurricane of controversy : the records tell us that it was denounced from every pulpit in Belfast. Yet, reading it now, one is impressed not only by its courage but by its reasonableness. It is true that he claimed that science will wrest from theology the entire domain of cosmological theory, whether of living or of non-living things ; but he had previously referred to the immovable basis of the religious sentiment in human nature, bringing as he said ' completeness and dignity to man.' The views of Lucretius and Bruno, Darwin and Spencer, might be wrong : whether right or wrong, he insisted, we claim the right to discuss them. If today I claim the right to discuss not only the scientific facts themselves but their consequences in human affairs, I doubt whether denunciation will follow : if it does, I shall be sorry but unrepentant.

The development which has brought most vividly to the public conscience today the ethical problems aroused by the advance of scientific knowledge lies in the field of nuclear physics ; and groups of scientific people in the free countries of the world are vigorously debating its various consequences, among them particularly the secrecy attached to weapons as new and devastating as those provided by nuclear fission. Atomic physics, however, is only one of many scientific developments which have brought, or are bringing, a mixture of possible good and evil about which judgments of relative value must be formed : we must not get too excited about one of them. There is no secrecy about most of these developments, they occur gradually and continuously before our eyes, we tend to accept them without question as though they were natural phenomena : yet in fact the consequences of one of them provide the most solemn problem in the world. The dilemma is this. All the impulses of decent humanity, all the dictates of religion and all the traditions of medicine insist that suffering should be relieved, curable disease cured, preventible disease prevented. The obligation is regarded as unconditional : it is not permitted to argue that the suffering is due to folly, that the children are not wanted, that the patient's family would be happier if he died. All that may be so ; but to accept it as a guide to action would lead to a degradation of standards of humanity by which civilisation would be permanently and indefinitely poorer. Conduct usually falls short of principles : but that would be the worst reason for abandoning principles altogether.

In many parts of the world advances in public health, improved sanitation, the avoidance of epidemics, the fighting of insect-borne disease, the lowering of infantile death rates and a prolongation of the span of life have led to a vast increase of population. Not only is the

population increasing but in many places its rate of increase is still rising : and these processes will take so long to reverse that for many years to come the shortage of natural resources, particularly of food, is bound to provide increasing deprivation and disturbance. That provides the practical motive of the Colombo Plan. In India, a Government Planning Commission in a report of July 1951 entitled ' The First Five Year Plan,' has dared to face the facts. A doubling in the last thirty years of the survival rate (births minus deaths) has led to a rate of increase of nearly 1½ per cent. per annum, a total of 5 millions every year in a population of 360 millions.

> ' With all the effort that the First Five Year Plan will represent, it will be possible barely to restore by 1955–1956 the pre-war standards in regard to food and clothing. Increasing pressure of population on natural resources retards economic progress and limits seriously the rate of extension of social services so essential to civilized existence.'

The pre-war standard in fact was miserably poor, a large part of the population existed below the level of a decent life, scores of millions only just above that of famine. Yet the gigantic national effort proposed in the Five Year Plan, even if successful, may only just restore that miserable standard. Can it sustain it then if the rate of population increase continues ? It is easy to answer that a higher standard of life has led in other countries to a gradually falling birth rate : but a higher standard requires a far greater charge on natural resources of all kinds, which cannot be met until the pressure of population is reduced.

In the meantime there is more than danger that the emergency will result in an over-use of natural resources, leading by land erosion, deforestation and other factors to permanent and irretrievable loss : this has happened already, and is visibly happening now, in many parts of the world. In a special Section on ' Family Planning ' the Indian Report recognises that ' an alteration in population trends takes at least a few generations to materialize; ' and steps are suggested for the education of public opinion on the need for limitation, and for experimental efforts to be made in the application of simple methods of birth control. For its wisdom and courage in acknowledging the gravity of the situation the Indian Planning Commission deserves every support : but the problem itself has not begun to be solved, and its consequences will dominate the development of India for many years ; indeed, its gravity will continue to increase. Malaria is admitted by the Planning Commission to take an annual toll of a million lives, tuberculosis of half a million. The resolute use of insecticides and anti-malarial drugs could soon reduce the former to a small fraction : tuberculosis is bound to require more effort and a longer term. Nobody would dare to say that steps to combat these diseases, and others such as cholera, to improve rural and industrial health, to increase the supply of drugs and medical equipment and services, should not be taken on the highest priority : but the consequence must be faced that a further increase of a million people per annum would result. Thus science, biological, medical, chemical and engineering, applied for motives of decent humanity entirely beyond reproach, with no objectionable secrecy, has led to a problem of the utmost public gravity which will require all the resources of science, humanity and statesmanship for its solution.

The example of India has been taken because of the sheer magnitude of the problem and because its seriousness is now admitted by humane and responsible men : but the same conditions exist already in many parts of the world and will soon exist elsewhere. It is *not* a question only of food : if a higher standard of life is to become universal, with education, communications, housing, reasonable amenities and public health, a far greater demand will be made on all such natural resources as power, chemicals, minerals, metals, water and wood. One is left wondering how long these can possibly take the strain. Could world supplies conceivably hold out if the present requirement per head, in the United States, were multiplied in proportion to meet the same demand everywhere—even without any increase of present population ; and if so, for how long ? There is much discussion of human rights. At what level

can these be reasonably pitched ? and do they extend to unlimited reproduction, with a consequent obligation falling on those more careful ? These problems must be faced not only with goodwill and humanity, but also with integrity and courage, not refusing to recognise the compulsion of simple arithmetic. It is right that the scientific imagination should be allowed to play sometimes with the more distant future, when possibly new methods and resources may be found to solve all problems ; but only on condition that our minds are not deflected from the urgent realities of the present.

By vast improvements in communications, which have made the world so small, applied science has been one of the chief agents in the present ferment of social, political and economic thought. Can one urge, after the event, that the application was a mistake and that the majority of mankind could better have remained isolated and in ignorance ? By making world war technically possible, applied science has helped to stir up national ambitions and social revolutions which, if poverty and deficiency continue without hope, may lead to major world catastrophe. Should we therefore refuse to employ science in defence of liberty and resign ourselves to a universal police state where no scruples are permitted ? Are we, in scientific research, to say that some subjects may be investigated, but not others for fear of the consequences ? Who then is to decide and by what international authority ? And is it practical to insist that all scientific knowledge should be fully and openly disclosed, without secrecy or reservation of any kind, military or industrial ? These are problems which cannot be solved by rhetoric, or by any simple formula. The purpose of setting them out is to make clear that we must face them with honesty and courage ; for they will not solve themselves.

I have led you to the ethical dilemma which perplexes many of us by taking an example in which few would question either the motives of those who made the original discoveries, or the humanity of their application : or indeed could wish that the fruit of the tree of knowledge had been left untried. It is easy to say now that side by side with the control of disease there should have been an equal and parallel effort in education, particularly the education of women as responsible citizens : for there is no possibility, if women remain ignorant and illiterate, of intelligent widespread family planning and control. But education alone would not have been enough, or indeed possible itself without a substantial measure of material and social betterment : and the expense and effort involved in this would have been indefinitely greater than in the application of medicine and hygiene, which after all has been relatively cheap. Had it been possible to foresee the enormous success of this application, would humane people have agreed that it could better have been held back, to keep in step with other parallel progress, so that development could be planned and orderly ? Some might say yes, taking the purely biological view that if men will breed like rabbits they must be allowed to die like rabbits, until gradually improving education and the demand for a higher standard of life teach them better. Most people would still say no. But suppose it were certain now that the pressure of increasing population, uncontrolled by disease, would lead not only to widespread exhaustion of the soil and of other capital resources but also to continuing and increasing international tension and disorder, making it hard for civilisation itself to survive : would the majority of humane and reasonable people then change their minds ? If ethical principles deny our right to do evil in order that good may come, are we justified in doing good when the foreseeable consequence is evil ?

I remember asking an eminent Indian who had taken part in drawing up the so-called Bombay Plan of 1944 why there was no mention of the gravest problem of all, overgrowing population : he replied that his colleagues and he had indeed discussed it, but decided to leave it to God. To a biologist aware of the methods by which animal population is in fact controlled by nature, this seemed pretty poor comfort : yet there are many who really take that view, admittedly with the element of reason that we never can be sure that things may not turn up to make all our calculations wrong. Should we then just continue to do the good we see in front of

us, in confidence that if our motives are humane, good and not evil will finally result? Or, taking that rather easy course, are we not showing a lack of the fundamental virtues of courage and integrity?

The dilemma is a real one, and cannot be resolved by any simple expedient. In another form it is perplexing many of those who are concerned with the development of nuclear physics, the ultimate service of which may be very great, possibly essential if our present type of civilisation is to continue when other sources of power dry up; while the benefits to medicine and industry are already substantial. *But*—nuclear fission has released the threat of unprecedented violence, with the possible destruction of many millions of lives and the accumulated treasures, moral and material, of civilisation. The individual conscience may tell a man to have no part in it: that is easy enough, for there are plenty of other interesting things to do, but it does not solve the problem. Moreover, it is possible that defensive weapons, based on nuclear fission, but not of the type intended for mass destruction, can be developed which would make armed aggression intolerably costly. What then of the abolition of secrecy? In principle, yes, for the historic and unique contribution of science to international goodwill has been in sharing knowledge regardless of race and frontier, and the chief satisfaction of scientific work, the condition of its fruitful development, is frank and free discussion. 'Cast thy bread upon the waters, for thou shalt find it after many days,' is wise and acceptable counsel in dealing with scientific knowledge: while 'he that observeth the wind shall not sow and he that regardeth the clouds shall not reap' is as aptly applied to human relations as to agriculture. Every possible endeavour, therefore, should be made towards international agreement on sharing scientific and technical knowledge and controlling nuclear weapons: but this, like peace itself, is a concern of every citizen, not only of scientific people. It is hard enough to get international agreement in quite simple matters, such as the perilous state of the north European fisheries, where no secrecy is involved and little national prestige, and the scientific evidence is unequivocal: but we must go on trying.

Much scientific and technical advance has led to unexpected dangers and difficulties. Without our present knowledge of bacteriology and preventive medicine, gigantic armies could never be kept in the field, and land war on the recent scale would be impossible: is medical science, therefore, to be blamed for twentieth-century war? The indiscriminate use of insecticides, by upsetting the balance of nature, can quickly do more harm than good. Radio communication may be used for spreading lies and disorder as well as truth and goodwill. Developments in microbiology, in many ways beneficent, may be used in the future for biological warfare, with effects at present unpredictable; and control by international agreement and inspection might be very difficult. The list need not be multiplied, all are aware that every new benefit to mankind provides also its own dangers, either as unexpected consequences or by deliberate misuse. Science is not alone in this: liberty may lead to licence, religion can be used to inflame passions, laws can be exploited to protect wrongdoing. If scientists feel called upon to examine their consciences, so much the better: but they need not imagine that in this they are exceptional!

It has been debated whether 'the scientific mind is fundamentally amoral.' The real answer is that there is no such a thing as '*the* scientific mind.' Scientists for the most part are quite ordinary folk. In their particular scientific jobs they have developed a habit of critical examination, but this does not save them from wishful thinking in ordinary affairs, or sometimes even from misrepresentation and falsehood when their emotions or prejudices are strongly enough moved. Their minds are no more amoral than those of surgeons, lawyers, or scholars. As investigators most of them realise that their function would be stultified were they to introduce moral data into a scientific argument. A surgeon is not required, or indeed allowed, to consider whether it would be better for the world if his patient died under the operation, he has only to carry it out with skill, care and integrity: but it would be

foolish to conclude that the surgical mind is amoral. The surgeon himself, as a human being, has to make moral judgments : but he does so outside the operating theatre. So it is with scientific people : like all good citizens they must take account of ethical considerations, and the chief of these, as with other good citizens, are of integrity, courage and goodwill. Integrity forbids them to allow feelings of any kind to obscure facts, but that does not make them amoral : after all, integrity is the first condition of morality.

In the practical world of today, complete abandonment of secrecy, in government and industry, is out of the question. The advantages to international relations, and to general scientific progress, of the greatest possible freedom are evident ; to these can be added the impossibility, in a free democracy, of keeping the best people unless the conditions of their work are congenial. If scientific men consistently avoid jobs which seem to them to fall short of reasonable freedom, they will force changes of organisation so that only necessary secrecy is maintained. The penalty of filling an organisation, governmental or industrial, with second-rate people, cheerfully amenable to unnecessary restrictions, is far too evident in its result on efficiency to be tolerable for long. The cure, therefore, is largely in scientists' own hands. In this, as in many other aspects of their work, moral considerations come in, and the only way to resolve the dilemma which is in so many minds is to discuss it frankly. To neglect it altogether is not amoral but immoral, it is the duty of all of us as citizens to consider the ethical basis of our work.

Today when the public importance of science and its popular esteem may turn some people's heads it is well that scientists should realise that the prestige of science is not their personal property, but a trust which they have an obligation to pass on uncompromised to their successors. The popularising of genuine science is an important public service, we should all be ready to take our part in it according to our powers : but to use the general prestige of science as a bait to attract attention to pronouncements on other topics, for example on politics or religion, is a disservice both to science and the public. As a citizen I need no more justification than any other citizen in saying what I like about such things : but I have no right to pose as a representative of science in discussing them. In the days when the representatives of religion claimed supernatural knowledge of the natural world it was necessary to insist, as Tyndall did, that the natural world belongs to science. Fortunately those days are past. If they now claim that the facts and trends of overpopulation are not what we say, we can argue about that as a scientific question : but if they insist that its consequences should be left to God, they must allow us as citizens to take the opposite view. If political pressure were applied in any way to force conformity to particular scientific theories, as happened in Germany and is happening now elsewhere, then one's right and indeed one's duty would be clear, alike as citizen and as scientist. But the nature of political institutions and the policies of political parties are not scientific questions and how I choose to vote has nothing to do with science. Indeed, curious as it may seem, the spectrum of political opinion of scientists in Britain is much the same as that of other similar groups : though the loudspeakers are generally tuned in to one particular wave-band.

Some seventeen hundred years ago there crystallised out, from many centuries of experience of the human problems and ethical necessities of medicine, the so-called Hippocratic Oath. The obligation of integrity and trust ; the insistent claim of suffering ; the care of mothers and children ; solicitude for the old and weak ; the sanctity of human life : these are as vital a part of medical tradition as its science and its art. Practice, it is true, may lag behind principles, but at least such failure is regarded as discreditable ; and it is hard to imagine any kind of civilisation in which the ethical principles of medicine were disregarded. That is one reason why the future possibility of biological warfare is particularly repugnant. Today science finds itself, unexpectedly and without those centuries of tradition and experience, in a position no less important to the community than medicine : and its ethical principles have not yet clearly emerged.

Every candidate for admission to the earliest of American learned associations was required to answer yes to the question, 'Do you love truth for truth's sake, and will you endeavour impartially to find . . . it for yourself and communicate it to others?' That affirmation might have its place in a modern scientific version of the Hippocratic Oath. But again the same dilemma arises—' endeavour impartially . . . to communicate it to others.' Apart altogether from considerations of national security, in many fields today much of the best research is done, and done increasingly, in industrial laboratories. Those who have seen and admired such work, and the people who do it, cannot but applaud the foresight which made it possible. But if all the results are to be communicated at once and impartially to others, could directors and shareholders be reasonably expected to continue their support? Indeed, if an industry were nationalised could it afford to give away its secrets to competitors abroad? Not in any real world, in which a nation must remain solvent and industry must depend for success on the rapid application of new knowledge. The dilemma must be met by reasonable compromise, of which perhaps the most hopeful sign today is that many of the directors of industry come up through research departments. A friendly and familiar contact between management and research, and between industrial and outside scientists, can reduce this particular dilemma to manageable size.

A graver problem is provided by research under government, when considerations of security come in. In the emergency which became evident in 1935, the secret development of radar for purposes of air defence aroused no obvious pangs of conscience; and many other developments come in that class. But the surest of military maxims is that counter attack is an essential part of defence; to limit scientific methods to defensive weapons would be to ensure defeat, indeed it is quite impracticable. But let us be realists; so long as offensive weapons may be used, the part played by the scientist is no more immoral than that of the engineer, the workman, the soldier or the statesman, and the attitude of 'holier-than-thou' is unbecoming. We all bear, as citizens, an equal responsibility. But is it practical to suggest that all scientists in all countries should agree, and hold to their agreement whatever happens, to take no part in research on offensive weapons? or at least should endeavour impartially to communicate its results to others? The answer is evident. There are individuals in all free countries who find such work intolerable. In those countries their scruples are respected and they are at liberty to do something else; but let them not imagine that the problem is solved that way, or that those who think otherwise are necessarily stupid or immoral. The first condition of freedom is freedom of conscience, and the scientist has the same right to that as any citizen: but freedom does not extend to giving away other people's property, whether of goods or knowledge.

There seems to be no simple answer to the riddle. All knowledge, not only that of the natural world, can be used for evil as well as good: and in all ages there continue to be people who think that its fruit should be forbidden. Does the future welfare, therefore, of mankind depend on a refusal of science and a more intensive study of the Sermon on the Mount? There are others who hold the contrary opinion, that more and more of science and its applications alone can bring prosperity and happiness to men. Both of these extreme views seem to me entirely wrong—though the second is the more perilous, as more likely to be commonly accepted. The so-called conflict between science and religion is usually about words, too often the words of their unbalanced advocates: the reality lies somewhere in between. 'Completeness and dignity,' to use Tyndall's phrase, are brought to man by three main channels, first by the religious sentiment and its embodiment in ethical principles, secondly by the influence of what is beautiful in nature, human personality or art, and thirdly by the pursuit of scientific truth and its resolute use in improving human life. Some suppose that religion and beauty are incompatible: others, that the aesthetic has no relation to the scientific sense: both seem to me just as mistaken as those who hold that the scientific and

the religious spirit are necessarily opposed. Co-operation is required, not conflict : for science can be used to express and apply the principles of ethics, and those principles themselves can guide the behaviour of scientific men : while the appreciation of what is good and beautiful can provide to both a vision of encouragement.

Is there really then any special ethical dilemma which we scientific men, as distinct from other people, have to meet ? I think not : unless it be to convince ourselves humbly that we are just like others in having moral issues to face. It is true that integrity of thought is the absolute condition of our work, and that judgments of value must never be allowed to deflect our judgments of fact. But in this we are not unique. It is true that scientific research has opened up the possibility of unprecedented good, or unlimited harm, for mankind ; but the use that is made of it depends in the end on the moral judgments of the whole community of men. It is totally impossible now to reverse the process of discovery : it will certainly go on. To help to guide its use aright is not a scientific dilemma, but the honourable and compelling duty of a good citizen.[1]

[1] Delivered Sept. 3, 1952.

SOME CONTRIBUTIONS TO EXPERIMENTAL PHYSICS FROM BRITAIN

Address by

PROF. A. M. TYNDALL, C.B.E., F.R.S.

PRESIDENT OF SECTION A

THOSE who are familiar with the works of Bernard Shaw will remember the character of Ann in *Man and Superman*. At one point in the drama Ann's mother, ineffective in parental influence and perplexed by the course of events, is driven to say 'It's a very queer world. . . . Nothing has been right since that speech that Professor Tyndall made at Belfast.' This reference to the address in 1874 by my namesake evoked a rich response from the patrons of the Court Theatre in London thirty years later, even if it passes almost unnoticed by an audience of today. I have no topic with which to follow John Tyndall in providing a landmark of controversy in the history of the Association ; still less am I likely to offer material for a shaft of humour from a future G.B.S. Indeed, it was with some hesitation that in retirement from academic life, I accepted the honour of the presidency of Section A. A. N. Whitehead once said that the aim of a university professor should be to exhibit himself as an ignorant man, thinking. I can only conclude that on occasions it is also the rôle of a professor emeritus, even though he no longer holds a chair in which to sit.

I must first make a brief reference to the death of an old friend of the Association and a valued supporter of Section A. Since 1920 Professor Allan Ferguson attended nearly every meeting of the Association until increasing ill health sapped his physical powers. He was its General Secretary from 1935 to 1946. In his Address as Section A President in 1935 at Blackpool, he gave evidence of the width of his philosophic and scientific interests. He was the embodiment of the spirit of the policy of the Association to act as a meeting ground between colleagues in sister sciences. Moreover, the records of the Association are full of the traces of that vigour, wit and humanity which were so well known to his friends and old students.

We have also to record our deep regret that Professor J. J. Nolan, whose name appeared in the preliminary programme as one of our Vice-Presidents, died last April. He and I shared a common interest in the study of gaseous ions, and he carried out a great deal of interesting work on condensation nuclei and ions in both town and country air. He often attended B.A. meetings in earlier days, and I had indeed looked forward to the renewal of personal contact with him that the war had broken.

It is now forty-eight years since I first attended a meeting of this Association early in my academic career. The 1904 meeting was at Cambridge, and it was my first visit to the Cavendish Laboratory. It has interested me to recall what I have remembered of that meeting, and even more so what I had forgotten. It must surely have been one of the most distinguished gatherings of physicists in the history of the Section ; yet, out of all their contributions, there are only three which have remained in any detail in my memory.

The first could almost stand as a testimonial to the value of visual aids in education—I remember it so clearly. It was a demonstration by R. W. Wood of the anomalous—as it was then called—dispersion of sodium vapour, familiar to many students of Physics as illustrated in the frontispiece of his text-book of Physical Optics. Wood was a constant visitor to

Great Britain in those days and clearly found an atmosphere congenial to him and his particular experimental genius. Starting with an empty bench and finishing only 30 minutes later with a ruin of devitrified glass, wet cotton waste and sputtering sodium metal, we had witnessed a demonstration which in simplicity, elegance, and slickness of manipulation, I have never seen excelled.

The second was a discussion on the so-called N rays claimed by Blondlot to have the properties of a weak beam of X-rays yet refractable by a prism. A show of hands disclosed a surprising number of eminent people who had tried, and failed, to repeat the results, subsequently proved to have been purely subjective in origin. The incident may have warned me early in my academic life of the trouble that can be caused by failure to make suitable experimental checks before rushing into print with sensational results. But the lesson should still be heeded today.

The third item was a paper by Kelvin who, with an inflexibility of opinion which is often the penalty of old age, advanced a molecular model for the emission of α-particles from polonium, as the last attempt to save the conception of an indivisible atom. This seems strange reading today when ' atomic energy ' is a term on every lip. But it must be remembered that Kelvin had belonged to a generation which hoped that by levers and screws it might pluck out the heart of the mystery of the Universe. Larmor used to tell the story of Boltzmann arriving, by a series of misfortunes penniless and without luggage, in Cambridge, whence he had travelled in search of Maxwell's model of the ether, a device of vortices and idle wheels which simulated the propagation of electromagnetic induction. If the model was not actually working, he at least hoped to find it in the cellars of the Cavendish Laboratory.

Indeed, Kelvin, had not been alone in his attitude towards the new ideas. Only the previous year Poynting, as President of Section A, speaking of ideas of atoms as storehouses of immense quantities of energy said, ' So wildly extravagant do they seem to many chemists and some physicists that they will not consider them with any patience. But the facts are extraordinary and had anyone described them ten years ago he would have been set down as fit only for an asylum.' But only two years earlier, as an undergraduate, I had been following Rutherford's early papers on radioactivity, at one stage twelve in one year from his pen alone in the *Philosophical Magazine*. Not only were they easy to follow, but they aroused such absorbing interest in me that I awaited the arrival of each monthly number as if for the next instalment of a serial story. So I listened to Kelvin on that occasion with some youthful impatience.

Since then, by frequent practice, we have become more receptive to unfamiliar conceptions involving new language of expression. Relativity, quantum theory, wave mechanics, meson fields have followed one another in quick succession. Today, through the discovery of a whole range of transient forms of matter, we seem to be entering yet another decisive period in this respect in the history of physics.

In the citation of a recent award of the Nobel prize to an English physicist, these words occur : ' His special claim to consideration is, in my view, the fact that he has shown that discoveries of fundamental importance can still be made with the simplest apparatus.' The speaker was referring to Professor C. F. Powell and his use of a photographic emulsion to record the tracks of charged particles of high energy. My reason for quoting from the occasion is that, in my view, the words emphasise an approach to the field of experiment which is characteristic of much of the best of British physics and in accordance with the great scientific tradition of the British race.

Examples of simple experiments which have opened up large fields for discovery are common in our literature. Each of us indeed might make our own selection. Immediately I think of Rayleigh in 1890 adding weighed amounts of olive oil to the water in a bath in order to find the least thickness of a surface layer which would impede the movement of camphor under surface tension forces. This simple experiment not only gave an estimate of molecular dimensions, but led the way to

the conception of a monomolecular film with all its applications to liquid and solid surface phenomena.

Or I recall Vernon Boys revolutionising a corner of instrument design by shooting an arrow across the floor of the laboratory, so that along it trailed a thin fibre drawn from a molten lump of quartz. Or my old teacher, A. P. Chattock, describing at the B.A. meeting of 1898 the first determinations to be made of the mobility of both positive and negative gaseous ions, by measurements of the pressure of the electric wind ; devising for it a sensitive pressure gauge which was taken by Stanton to the newly opened National Physical Laboratory in 1901, and used there in the first accurate and reliable determinations of the pressure of wind on structures, and in all the early pioneering wind-tunnel investigations in aeronautics.

But instead of building a catalogue in this way, let us turn to the programme of the present meeting and trace some of its items back to their early days.

The history of electrical discharge in gases, the subject of one of our discussions, is a long one. Robert Boyle was detecting the chemical action of electrical discharge when he noticed the smell that accompanied the rubbing of amber. But the subject will always be particularly associated with the work of J. J. Thomson, Townsend and Rutherford in Cambridge over fifty years ago, when by simple means they explored the phenomena of gaseous conduction in ionisation chambers. The subsequent ramifications of these experiments have been legion. Yet that period has been termed the age of sealing-wax and string ; when with the eyes of the whole scientific world focused on the Cavendish Laboratory, its equipment in apparatus was probably inferior to that of a number of its continental competitors.

Today we are still concerned with the phenomena of ionic mobility, recombination and diffusion, then being investigated for the first time. Modern work on the electrical breakdown of gases by disruptive discharge which appears in our programme is an outcome of the early experiments by Townsend in which he carried the voltage-current characteristic of an ionisation chamber into the region where ionisation by collision between electrons and molecules sets in. Nuclear physics sees another application of this early work in the proportional counter used, for example, with good effect recently in Glasgow by Curran and his collaborators in the study of soft radiations and slow electrons.

The same period saw the interpretation of ionic diffusion in terms of classical kinetic theory, modified a few years later by Langevin to allow for the effect of the charge on the ion in creating induced electric dipoles on the surrounding molecules, treated as elastic spheres in thermal agitation. Though the limitations of this early treatment are now apparent there are some collision phenomena for which it still gives a good working model. For instance we showed at Bristol that ions of the alkali elements move through pure gases at speeds agreeing in the main with this theory. Moreover when the attractive field is great enough it binds a cluster of molecules to the ion, despite the shocks of thermal agitation.

Thus a small ion like Li^+ will collect a cluster of highly polarisable atoms of xenon even at room temperature ; but, in neon, it is necessary to reduce the thermal velocity by working at 90° K before this effect is observed. In helium the effects are not appreciable until temperatures of the order of 20° K are reached. In the presence of traces of a highly polar impurity such as water vapour, the formation of clustered ions may be followed experimentally at ordinary temperatures. Indeed, without special precautions to avoid the presence of polar molecules, ions in the lower atmosphere normally exist in the clustered state.

But this simple classical picture is not adequate when we consider collisions between electrons and molecules, or between charged and uncharged molecules of the same gas. So here we have a link between the classical and the modern picture of events. The application of wave mechanics to a variety of collision and discharge phenomena is a field of work with which the name of Professor D. R. Bates of Queen's University is now closely associated.

The subject of the physics of the high atmosphere, also under discussion this week, affords a striking example of my theme. I have in mind the first experiment

of Sir Edward Appleton by which, with Barnett, he directly established the existence of the Kennelly Heaviside layer of ionised air. Yet it was merely the radio analogue of an experiment familiar to every student of Physics, namely, the production of Lloyd's fringes by interference between two trains of light waves from a line source, one of which in transit had suffered reflection at a mirror surface. In the optical experiment the pattern is observed as a variation in intensity with position. In the radio experiment it was more convenient to fix the receiver and vary the wave-length of the radiation. The waves were received in Oxford from a transmitter in Bournemouth and recorded on a fairly quick recording galvanometer. It must have been exciting to witness the indicator go up and down as the two sets of waves arrived in and out of phase, and to work out for the first time the height of the layer.

But this simple start by Appleton on the study of the ionosphere had an even greater significance. For one thing, it concentrated attention on echoes and so led to radar and the science of radio-meteorology. In turn that led to the detection of intense radio sources in outer space and so to a new science of radio-astronomy.

Radio-astronomy would have afforded me with good examples of pioneering work still in progress in Cambridge by Ryle and others and at Jodrell Bank, near Manchester, by Lovell and his collaborators. But as Professor Lovell has recently delivered in Belfast the Norman Lockyer lecture of this Association which he devoted to this subject, I will only briefly mention two items. Firstly, another radio analogue, in this case of Michelson's stellar interferometer, has been developed by Ryle to map the areas of far distant radio sources ; secondly, the latest news from Manchester is of the proposed start on the erection at Jodrell Bank of a new great radio-telescope, which should do for radio-astronomy what the large visual telescopes in the States have done for classical astronomy.

The next subject to which I wish to refer is the physical properties of materials in the solid state. Though the subject only enters into our discussions in a minor way, it is one which is being actively pursued on both sides of the Atlantic at the present time. It was Sorby in England who first subjected the surface of a solid opaque material to microscopic examination. But anyone who reads some later papers in 1900 by Ewing and Rosenhain on the microscopic study of the crystalline structure of metals will realise how much arose from them. Not of least importance among their observations was the discovery of the phenomenon of slip at gliding planes when plasticity starts.

The study of the solid state also owes much to the methods of X-ray analysis. Although the Braggs, father and son, were not the first in the field in the discovery of X-ray diffraction, their measurement of an atomic spacing by X-ray reflection had the kind of simple directness to which I have specially referred. Through it they founded the science of X-ray spectroscopy with its countless applications in both pure and applied science.

As one illustration of the present activity in Britain in the field of the solid state, we may take some recent work on the growth of crystals. For a long time physicists have been puzzled why crystals grow as fast as they do from supersaturated surroundings. Since its binding energy to a surface will depend upon the number of close neighbours, an atom is less likely to condense upon an ideally flat surface than at other places. Calculations indeed suggest that if growth depended upon this process, it is most improbable that the known rates could be reached.

On the other hand, X-ray diffraction studies have shown that most real crystals are imperfect. One common defect is a lattice misfit known as a dislocation, a conception with which the name of G. I. Taylor is closely associated, and which is now being extensively applied to explain some of the mechanical behaviour of crystalline solids. Now a principal crystal face which contains the end of a screw dislocation will have a step, one end of which terminates at the end of the dislocation. Such a step forms a much more favourable place for the retention of arriving atoms. But in this case it is easy to show that new layers never start, because condensation at the step does not eliminate the step. As the step advances

it acquires a spiral form, giving in the case of one dislocation a single layer in the form of a helicoid, instead of a series of parallel layers on top of one another.

When this theory was first put forward by F. C. Frank in 1949 at a conference in Bristol, it received dramatic confirmation by the production from the pocket of a member of the audience (L. J. Griffin) of a micro-photograph of this phenomenon on the surface of natural beryl ; a striking example of the interplay of theory and experiment. Since then many spiral patterns on crystals of compounds and of metals have been found, and the process of growth has been watched under the microscope.

Again, Dr. Keith, speaking later this morning, will be describing work that Dr. J. W. Mitchell and he have been conducting on the crystal forms of silver, produced in the reduction of silver bromide by standard chemical developers. What they have done is to produce a single crystal of pure silver bromide in the form of a laminar sheet about 0.2 mm. in thickness and of given orientation. They then study under the microscope the growth of silver crystals in its surface. They have in this way shown that even in the long-investigated subject of photographic action, an experiment framed on a simple basic principle can lead to a new theory of development.

Perhaps the best illustration of all comes from the field of nuclear physics. Much has already been written elsewhere on our past contribution to the foundations of the subject. Less has been said of Britain's present contribution, although, in one aspect of the subject resulting from the study of cosmic rays, it still leads the world. Moreover, in the techniques of detection of particles, whether by the trigger methods of scintillations or ionisation counters, or by the track methods of cloud chamber or photographic plate, Britain has from the start been a pioneer.

To anyone who wishes to find modern examples of photographs of tracks in a Wilson cloud chamber, I would refer them to a recent publication by Rochester and Wilson [1] which collects in one volume some 120 cosmic ray events. Some of the most novel of these refer to pairs of tracks forming a V which appear to originate in the body of the gas. They were first discovered by the Manchester group operating at the Jungfrau Observatory and their existence has since been confirmed in Pasadena. Momentum considerations show that there must have been some incident particle, (uncharged because it left no track), provisionally called a V particle, which decayed in flight into two charged particles, whose tracks form the V.

In a later Table, details concerning three kinds of transient V particles, two uncharged and one charged, are given, together with their probable modes of decay and an estimate of their average lifetime.

I myself cannot speak with any direct experience of this technique. But I have been privileged to watch the progress of the photographic emulsion technique perfected by Professor Powell in the Wills Laboratory at Bristol. As an example of the striking consequences of following up a simple idea there can be few better examples in the history of physics. The primary purpose of the cyclotron and its successors is to create artificially particles of high energy and to study the consequences of collisions between them and atomic nuclei. Machines are now under construction which will produce energies up to about 6×10^9 electron volts (6 BeV). But by suspending a box of plates from a balloon at $\sim$ 85,000 feet the properties of charged particles of far higher energies, up to 12,000 BeV, are being studied, as primary cosmic rays from outer space. Fundamental nuclear processes are thus under observation which are, and may always be, beyond the range of any laboratory machine. Moreover, these rich dividends come from an exceedingly modest expenditure, running as well as capital, compared with that normally associated with a large synchrotron.

It is interesting to recall that measurements on α-ray tracks in emulsions were being made in Rutherford's laboratory in Manchester as early as 1910, before the invention of the cloud chamber. But the results were quantitatively not promising, a conclusion again reached as late as 1935. However, in 1938, Powell showed that,

[1] Rochester and Wilson, *Cloud Chamber Photographs* (Pergamon Press).

even with the emulsions available at that time, the method had great potentialities. Later, thanks to the co-operative investigations of the photographic industry in England, standardised plates sensitive enough to record tracks of minimum ionisation, and with emulsions as thick as 600 μ were made available; and the method is now in world-wide use.

Some of the principles of measurement on the tracks are taken over from those of a cloud chamber. Provided that the particle is not moving at relativistic speed, the grain density may be used to estimate the speed of the particle at any point in the track. The particle also experiences small angle scattering as it shoots through the assembly of atoms in the emulsion; from this we may deduce its momentum at any point. These measurements may thus be combined to eliminate the velocity and to give the rest mass of the particle. An additional check is possible when a track ends in the emulsion, since the length of the track is a measure of the initial kinetic energy of the particle. The quantities Mv^2, Mv and v are then all known.

To identify particles of charge greater than unity another measurement is necessary, because grain counting in sensitive plates is impossible when the track becomes continuous. This is furnished by a study of the density of delta rays, which, as electron tracks, appear like hairs on the main track. Alternatively, insensitive plates have recently been used so that grain counting can be undertaken. A paper on the analysis of the primary cosmic rays by these methods by P. H. Fowler appears later in the programme.

Many relativistic particles giving thin tracks arise in high energy distintegrations; in some cases the tracks have been long enough to permit a mass determination to be made as the particle slows up.

It must be appreciated that the accuracy of such mass determinations is not high judged on the standard of mass spectroscopy. Deduced from a single track the possible error may be as high as 15 per cent. But, even so, the evidence for several new types of meson, included in the Table together with information on their properties, is strong.

The μ-meson of mass 209 electron masses known before the war was regarded then as confirming Yukawa's theory. In this theory Yukawa concluded that a particle of a transient type could arise in a collision of sufficient speed between a nucleon and a nucleus. The first new type to be discovered by this technique was the

Particle	Symbol	Discovery	Mass	Average lifetime in secs.	Remarks
Electron	e^-	Cambridge 1897	1	Stable	Can combine to give a γ-ray pair.
	e^+	Pasadena 1932	1	Stable	
Neutrino	ν°		<1	Never directly observed	
Mu-meson	μ^+	Pasadena 1936–38	209	$2{\cdot}2 \times 10^{-6}$	$\mu^+ \rightarrow e^+ + 2\nu^\circ$
	μ^-	„ „	209	$2{\cdot}2 \times 10^{-6}$	$\mu^- \rightarrow e^- + 2\nu^\circ$
pi-meson	π^+	Bristol 1947	276	2×10^{-8}	$\pi^+ \rightarrow \mu^+ + \nu^\circ$
	π^-	„ „	276	2×10^{-8}	$\pi^- \rightarrow \mu^- + \nu^\circ$
	π°	Berkeley 1950	262	$<10^{-14}$	$\pi^\circ \rightarrow 2h\gamma$. Occasionally $\pi^\circ \rightarrow \gamma + \beta^+ + \beta^-$
zeta-meson	ζ	Bristol 1951	535	$>10^{-11}$	$\zeta^\pm = \pi^\circ + \pi^\pm$
	ζ°	„ 1952	550	$<10^{-14}$	$\zeta^\circ \rightarrow \pi^+ + \pi^-$
Neutral V	V_2°	Manchester 1947	~800	$\geqslant 10^{-9}$	$V_2^\circ \rightarrow \pi^+ + \pi^- + ?$
tau-meson	τ	Bristol 1949	970	$\sim 10^{-9}$	$\tau^\pm \rightarrow \pi^\pm + \pi^+ + \pi^-$
Charged V	$V^\pm$	Manchester 1947	~1000	$\sim 10^{-9}$	$V^\pm \rightarrow ?^+ + ?^\circ$ Possibly identical with a kappa meson
kappa-meson	κ	Bristol 1951	1079	$\geqslant 10^{-9}$	$\kappa \rightarrow \mu + ?^\circ + ?^\circ$
chi-meson	χ	„ „	1450	$>10^{-9}$	$\chi \rightarrow \pi + ?^\circ$
Proton	P^+	Manchester 1911	1836	Stable	
Neutron	N°	Cambridge 1932	1836	12 mins.	$N^\circ \rightarrow P^+ + e^- + \nu^\circ$
Neutral V	V_1°	Manchester 1951	~2200	$\sim 10^{-9}$	$V_1^\circ \rightarrow P^+ + \pi^-$

π-meson of mass 276,[1] found at the Pic du Midi observatory (altitude 9,000 feet), and produced by a collision between a nucleus and an incident proton in cosmic rays. This is a particle which interacts strongly with nuclei, disintegrating any with which it may collide, but transforming into a μ-meson and a neutral particle at the end of its free life. On the other hand, the μ-meson, owing to its weak interaction with nuclei, can pass freely by them. These facts were disturbing to the Yukawa theory which had no place for a weakly interacting meson, nor, indeed, for a whole range of transient particles.

Reference to the modes of decay given later will show how important these two mesons are in the general scheme. Indeed, the discovery of the $\pi \rightarrow \mu \rightarrow e$ transformation, historically the first, was also probably the most important in its influence on the research. Firstly, it brought into relief the special characteristics of different types of track, and the advantage of having a medium of high stopping power in the study of transient charged particles. Secondly, it provided data and experience used later in the study of many types of nuclear ' stars ' and in the detection and properties of the still heavier particles. Thirdly, a photograph of this event gives indirect though unequivocal evidence of the existence of the neutrino, a neutral particle of small mass, postulated before the war as an accompaniment to β decay, but never isolated. From the principle of conservation of momentum, it is obvious that there must be a neutral particle ejected at the end of a π-track to balance the μ-meson shooting off from the same spot. Owing to the range of velocities of electrons from the $\mu \rightarrow e$ decay, at least two neutrinos are necessary in that transformation. Yet again, the result of introducing also the principle of conservation of mass energy, illustrates how beautifully the stages of the event hang together to show that the rest mass of the neutrino is less than twice that of an electron and may even be zero.

When nuclear disintegrations are caused by protons of relatively low energy, most of the fast secondary particles other than protons are π-mesons. At the extreme altitudes reached in balloon flights, high energy components of cosmic rays produce disintegrations in which the whole nucleus is violently disturbed. Accompanying the many nucleons evaporated in such processes are heavier mesons of several other types. The masses of these new bodies are still not known with any precision, but there is evidence of at least four particles, zeta-, tau-, kappa- and chi-mesons having separate identities and masses of the order of 530, 970, 1070 and 1450 respectively. Of these kappa- and chi-mesons have not been observed when the incident particle has an energy less than $1{\cdot}5 \times 10^{10}$ eV (15 BeV). Even this is far beyond the range available in the laboratory, so that it is possible that they will not be produced artificially, as π-mesons have been, for for some time to come.

In the Table of mesons and V particles with their estimated masses, lifetimes and modes of decay, the proton, neutron and electron have been added for completeness. At first sight this table appears to present a picture of considerable complexity. But closer examination shows that some order is beginning to emerge. Thus the constant appearance of the π-meson as a product of decay is significant. The zeta-meson charged or neutral, appears to decay into two π-mesons ; the tau-meson into three π-mesons. The chi-meson decays into a π-meson, and a heavy neutral particle. I am informed that a search is now being made for the decay products of this neutral particle. Tentatively, it is suggested that this may be a neutral tau-meson giving the reaction

$$\chi^{\pm} \rightarrow \pi^{\pm} + \tau^{\circ} \rightarrow \pi^{\pm} + \pi^{+} + \pi^{-} + \pi^{\circ}$$

There are still some outstanding difficulties, some of which may possibly be removed when more accurate data are available. It is, for instance, possible that a particle designated V is in some cases identical with one named differently by the Bristol group.

In any case, a wealth of new information is arising from a study of these collisions of great energy, and some remarkable phenomena are now being recorded. This is not the occasion to go into detail, but one example may be given in which an incident particle of energy about 2×10^{13} eV collides with a nucleus in a box of plates

[1] Later π-mesons were produced artificially in the Berkeley cyclo-synchrotron.

at 85,000 feet. A widely spread star of tracks of 22 heavy nuclei so produced is accompanied by a narrow jet of 76 particles following the direction of the primary. This jet can be followed through twenty successive emulsions in its passage through the whole box, with a secondary, tertiary and quaternary star *en route*, each initiated by a jet particle from the previous star. There is also evidence of the presence of high-speed neutral particles in the body of the jet.

We have thus come to realise that the stable forms of matter, protons and electrons which make up our familiar world, are only a small part of the variety of forms that matter can take. In the search for a theoretical explanation of the processes, the clues may already exist unrecognised in present data, or more surprises may yet be in store in the study of these events. Meanwhile the facts accumulate.

I recall that when I was recorder of Section A twenty-three years ago, I asked Lord Rutherford if he would give a general talk on the trends of modern physics. It was at a time when theorists were using the new ideas of wave mechanics with great effect. He refused, with the remark 'There is nothing to say, except that the theoretical physicists have got their tails up and it is time that we experimenters took them down.' It would have pleased him to find that today in the field that he did so much to create, the tables are turned, at least for a spell.

In concluding this survey, it is gratifying to recall that of the eight to ten transient new particles which have been discovered since the war, all but one have been found by experimental workers in the laboratories of Great Britain.

In stressing the fruitfulness of our traditional national approach to problems in experimental physics, it is not my wish to exaggerate. Science knows no boundaries, or should not; and many nations contribute, each by methods which may in their own way be equally traditional. But I believe that we can still take great pride in our continuing contributions to the subject, and there is something to be said for emphasising them at a time when in some other respects we no longer hold the position that we did in world affairs.

On the other hand, I have little belief in the importance of trying to establish claims to priority whether national or individual. It is often the circumstances in which a discovery is made that determine whether any use is made of it; two independent discoveries of the same fact may therefore suffer very different fates. So, in mentioning individual names, I have just thought of them as men who by simple means have had a profound influence as pioneering leaders in a given field to which others have also contributed.

But having said this, I also feel that, in changing circumstances, it is important for us to preserve conditions of work and training favourable to the development of these national qualities in those at the threshold of their careers.

Fifty or more years ago the manipulative versatility of many of our physicists often evoked surprised comment from continental visitors more dependent upon aid from technicians than we were. Many of us learnt with varying degrees of success to be our own carpenter, metal worker and glassblower in the simpler tasks. But today technicians are frequently available and material equipment more lavishly supplied to speed the progress of research. On the other hand, at the undergraduate level the introduction into the subject of new principles—relativity, quantum theory, wave mechanics, and the like—has increased the scope of the theoretical background necessary to a physicist, and has presented problems to the teacher in preserving a proper balance between theory and experiment without damping the ardour of the practically minded young man.

It would be a platitude to say that training in the subject should be such as to stifle neither an Einstein nor a Faraday: and I do not feel that any danger lies there. Some time ago a friend said to me 'We don't train a Vernon Boys nowadays'; one who was said to have personally worked and fashioned every material known to man. Here I think the answer is that we never *did*. To such a person the door of a university is still wide open and he will find no conditions

which he cannot surmount in the exercise of his genius.

The group that I have in mind is a much larger one, but at the same time containing many of promise in constructive ability. In my young days few of these ever entered the profession of physicist, because neither industry nor the civil service employed them; and university and even school appointments were relatively few in number. I refer to the group that mostly remain in the university for an additional two or three years in pursuit of a Ph.D. degree. Nowadays the success of many a research project, particularly in these days of increasing team-work, will depend enormously on the degree of initiative and resource that these men can show, when closely concerned with the details of the work. Indeed, there may also be others apparently less able, in whom dormant faculties may be awakened by stimulus coming from without, and to which in my experience they are often specially sensitive.

For these men the type of experience that they acquire in the Ph.D. course may be of vital consequence. In most cases at the time of embarking upon it, only three years have elapsed since sixth form work at school. If, therefore, in prosecuting research a student spends much of his time in collecting new data by an established technique with standard equipment, he may publish papers and write an impeccable informative Ph.D. thesis and yet fail in the end to be a trained physicist in the best sense of the word. Though one has some sympathy with his supervisor in using him as a research assistant on a project profitable in scientific results, it should be remembered in the national interest that slickness in a particular technique does not necessarily raise a man above the skilled artisan in his ultimate usefulness, to be paid for in wages from an industrial or a public purse.

It is obvious that all possible technical assistance should be given to fully trained men, so that their work can progress at maximum speed. On the other hand, I would plead that young post-graduate students, if no longer dependent upon their own efforts at the lathe or blowpipe, should still be so trained as to be able to design an apparatus on sound mechanical lines and follow it intelligently through the various stages of workshop practice.

It is sometimes easier to buy a new piece of expensive apparatus than to think. It was from the presidential chair of a previous meeting that I learnt of the charwoman, who, complimented by her mistress on the clever way in which she had surmounted some difficulty, said 'Well, it's like this, Mum, if you 'aven't 'ad education, you've just *got* to use your brains.' So a spur to the development of ingenuity should be preserved for the young physicist in these formative years.

It is, on the other hand, true that the modern range of technical knowledge in either theory or experiment calls for a degree of specialisation not necessary to Stokes, Maxwell or Rayleigh, whose equal facility in either sphere we are not likely to repeat today. For that reason the recent growth of theoretical research schools in close association with experimental teams is to be wholeheartedly welcomed. Judging from the titles of their posts, it seems a matter of accident whether these men are called applied mathematicians or theoretical physicists. Some would say that in any case they are one and the same, except that the former work on hydrodynamics and the latter on solid or nuclear physics. My own view is that a true association of theorists and experimenters in close contact gives to the theorist a method of approach which is physical in conception rather than mathematical.

Horace Lamb in Cambridge in 1904 spoke of his delight in watching Stokes' optical demonstrations, but added that at the time he thought it a little unnecessary to show practically that the height of the lecture room could be measured by the barometer, or to verify the calculated period of oscillation of water in a tank by actually timing the waves. I would say that this was not the standpoint of a physicist but of a mathematician. Still less would I wish to claim for my profession a young man of my acquaintance who, invited to acquaint himself with the contents of an elementary laboratory course, criticised the inclusion of the simple pendulum experiment. He said that he had looked up Kaye and

Laby's tables and found that the value of g was 981·17. He could not see how the student could be expected to improve on this value with the apparatus available !

Finally, it must be accepted that some measure of central planning of research has come to stay, if only with a view to preventing duplication in fields of heavy expenditure. But the inception of a new research is usually the effort of an individual even though it may lead to team work. It may indeed arise as a bright idea in the mind of a young man not yet established in reputation. Its ultimate direction will never be inspired unless it is in the hands of men following their noses on the job. Indeed, its very progress may depend upon some observation of stray occurrence made in this way. I am therefore conscious of the dangers of central financial control of research projects by panels of experts in given fields. It must be realised that they will not be supermen but ordinary human beings meeting in an office, having their own, and may be preconceived, views on what is important. For instance, from opinions expressed to me at the time, I am confident that no such body would have voted funds from the public purse in 1938 to enable a newcomer into the field of nuclear physics to reopen the study of particle tracks in emulsions.

Again, the establishment of special research institutes each confined to certain fields of work no doubt has certain advantages ; a project should not thus suffer from lack of resources. Indeed, research laboratories associated with particular industries are naturally organised on this basis. But in the one laboratory that I have directed, I always resisted the temptation to concentrate the resources in one field, and I believe that the Wills Laboratory in Bristol has in no way suffered from the versatility of its interests.

Rutherford once explained in a letter to the *Philosophical Magazine* how as a young man he was led to investigate the ionisation method of studying radioactive radiations. It arose, he said, because at Cambridge the supply of radioactive material was too meagre for the standard method of photographic blackening to be applied. Being Rutherford, had he gone to Paris, where the necessary facilities were to be found, he might still have blazed a trail, even though a different one and not in Cambridge. But the story emphasises an aspect of research which any co-ordinating plan must at all costs safeguard, and it will have to be an intelligent one to do so.

In speaking of our national tradition in experimental physics, I have indicated certain pitfalls to be avoided, if it is to be maintained. The details of any practical policy must be the responsibility of younger men. Indeed, the leaders in any field of scientific exploration must always have the vigour and initiative of relative youth. In them I put my trust, provided that they feel that they also can learn something from the past that I remember.[1]

[1] Delivered Sept. 4, 1952.

THE ADVANCING FRONT OF CHEMISTRY

Address by

PROF. W. WARDLAW, C.B.E.

PRESIDENT OF SECTION B

SOME wise man once said that men more frequently need to be reminded than informed. I do not hope or expect today to tell you what you do not know or to speak of anything outside the range of reflections which at one time or another have been familiar to most of us. Further, I am fully mindful that a Presidential Address must, from its very nature, present a mere outline of a topic.

In his *British History of the Nineteenth Century* G. M. Trevelyan has stated that :

> 'During the last 150 years the rate of progress in man's command over nature has been ten times as fast as in the period between Caesar and Napoleon and one hundred times as fast as in the slow prehistoric ages. Tens of thousands of years divide man's first use of fire from his first application of it to iron . . . the art of writing preceded the printing press by tens of centuries. In those days each great invention was granted a lease of many ages in which to foster its own characteristic civilisation before it was submerged by the next. But in our day, inventions, each implying a revolution in the habits of man, follow each other as thick as the falling leaves.'

Briefly, the pace of scientific, social and economic changes in modern times is unprecedentedly rapid. It is easy to understand why the pace of our science of chemistry is growing ever faster and faster. In the last quarter of a century the number of investigators in both pure and applied chemistry has enormously increased, whilst research itself is more adequately financed and to some extent better organised than ever before. It is tempting but not very profitable to try and imagine what will have been accomplished, say fifty years hence, if this acceleration continues. Many have indulged in speculations on the intellectual reactions which the great pace of the advancing front of science may involve. It has been suggested that life may not be long enough for the would-be investigator so to store his mind as to reach the growing frontier where his work should begin. Again, chemists may well consider that the pace of advance may be such that no human mind will be able to correlate into any significant whole the enormous accumulation of chemical facts. Yet I believe that man's curiosity about the universe and his own complex self will increase and that the astonishing growth of our science of chemistry will continue.

Chemistry has been defined as the science of substances—their structure, their properties and the reactions that change them into other substances. For a long time the story of chemistry was a record of the study of up to 92 chemical elements and their compounds. The 26 letters of the alphabet can be combined into hundreds of thousands of words each with a different meaning. What then are the potentialities before the chemist with some 92 elements at his command? Nevertheless, at intervals during the last half-century prophets have confidently predicted the demise of inorganic chemistry and even of organic chemistry. Chemistry however, has confounded the prophets by its amazing vitality. The award of the Nobel Prize for chemistry for 1951 to E. M. McMillan and T. Seaborg directed the attention of the whole scientific world to the most spectacular advance made in chemistry for a long time : the extension of the Periodic System from element 92, uranium, to element 98, californium. Elements heavier than uranium could

not be found in nature and speculations about their possible existence were freely made. In 1934 Enrico Fermi and his collaborators in Rome placed a small piece of radium and beryllium against some oxide of uranium. Some of the neutrons emitted from the radium-beryllium source were lodged in nuclei of uranium atoms and as the new form of uranium had a β ray emission Fermi thought he had produced an element higher up in the Periodic Table. These researches were continued in Paris and Berlin and for several years it was believed that four or five successive transuranium elements could be synthesised by this simple method. In 1938 and 1939 the work of O. Hahn, F. Strassmann and Lise Meitner disclosed that scientists had not been creating new elements after all—instead they had been splitting uranium into two parts. The first real transuranium element, number 93, was discovered by E. M. McMillan and P. Abelson early in 1940 and from then on, progress was rapid. With the help of powerful cyclotrons and the newly constructed uranium piles, a large team of workers concentrated on the task of producing new elements and Seaborg and his co-workers have identified elements 93, 94, 95, 96, 97 and 98, and much has been learned of the chemistry of these transuranium elements. It is natural to ask, will other transuranium elements be identified and the table extended beyond 98? Drs. Seaborg and McMillan think so. Further, they feel they can predict the chemical properties of these elements of the future. Particularly important in this work was the development of ultra-microchemical methods of analysis which enabled the chemistry of the transuranium elements, usually available in millionths of a gram, to be investigated. The combined discoveries of the physicist McMillan and the chemist Seaborg and their associates have opened whole new fields for the investigator interested in the realm of inorganic chemistry. These results were made possible by a type of research organisation which nowhere has been brought to such perfection as in the big scientific laboratories of the United States. Ample means were placed at the disposal of the investigators in view of the great practical importance of the results to the atomic energy project.

Chemistry already has made an indispensable contribution to atomic energy research and development and chemists will fill a major rôle in establishing nuclear power as a benefit to humanity. The next phase in the development of atomic power probably depends more on new materials than on any discovery in nuclear physics or improvement in atomic technology. This extension of technological interest in new metals, all over the world, is bound to emphasise the need for broadening our knowledge of both the familiar and unfamiliar elements and compounds. It is surprising how little we know about many of the chemical elements, even the metals. We know still less about certain elements which lie midway between the metals and the non-metals, the so-called metalloids. A host of new researches have been set in motion primarily by the needs of the atomic energy programme. However, the results may well be the starting-points of new developments in every phase of industry.

In the U.S.A. there is a great revival of interest in inorganic chemistry and in this country there are signs that its value and importance are being increasingly realised. The development of inorganic chemistry has been handicapped in the past by a lack of appreciation of its content and a misunderstanding of the place of technique in science. The essence of the several branches of chemistry lies in their objectives and not in their techniques. Whilst the content of organic chemistry as the chemistry of carbon is generally accepted, many who should know better, regard the distinction between physical and inorganic chemistry as implying the distinction between a quantitative and a descriptive approach. I take the essential character of physical chemistry to be the investigation of general laws and universal phenomena, whilst inorganic chemistry is the study of all that is specific to each of the elements and their several combinations each with one another. If one may judge by his new book *The Structure of Physical Chemistry*, Sir Cyril Hinshelwood shares this view-point, for he does not introduce particular chemical species or reaction mechanisms into the discussion

until the latest possible stage. It is evident that the same problem can frequently be regarded either as physical or inorganic chemistry according to the aim in view. The kinetics of the hydrogen-halogen reaction is properly described as physical chemistry if the interest lies primarily in general reaction mechanisms and kinetic and statistical mechanical theory. If the interest lies in comparing the energetics and specific reaction mechanisms of the different halogens it is essentially inorganic chemistry. From this standpoint there is more inorganic work being done in British Universities than would appear at first sight. However, greater interest will undoubtedly be kindled in this branch of our science by the march of events in the industrial world.

The contribution of chemistry to the development of atomic energy is not always fully appreciated. Because of the physical principles involved it is sometimes said that atomic power is principally the concern of physicists. Obviously anyone concerned with this field of work should be familiar with the basic facts of nuclear physics. However, it may be stated quite fairly that the economic future of atomic energy calls for the close association of scientists and engineers and that the rôle of the chemist is a vital one. Amongst the problems which arise in inorganic and physical chemistry are those concerned with the behaviour of solid and liquid inorganic salts and metals at high temperatures. Of primary interest is the solution chemistry of many inorganic ions and complexes, whilst the effect of radiation on various kinds of liquids, solids and solutions poses problems of great complexity. Basic studies of corrosion phenomena are of considerable importance, whilst the search for new materials for constructional purposes is imperative. Most of the elements have been surveyed for their neutron absorption and as a result special attention has been focused on the chemistry and metallurgy of beryllium, zirconium and other metals and alloys. It will be apparent that in pressing ahead for practical results it has been necessary to forego much fundamental work. In due course it will be desirable to fill in the scientific background. These few remarks indicate the range of problems confronting the chemist who finds in the advancement of knowledge an engrossing adventure.

A question of great general interest at the present time is how far are we entitled to visualise the peaceful development of atomic energy in the twentieth century. Scientists and engineers have expressed their views on this matter, but a recent article in *The Three Banks Review* by Professor R. S. Sayers, Professor of Economics in the University of London, calls for special study. He expresses the opinion that the possibilities of peaceful exploitation are not fantastically speculative if we confine our view to the next generation or so. For example, if coal prices rise relative to engineering costs the date at which atomic furnaces are an economical proposition is brought nearer. Secondly, the cost of constructing plant of this new kind is highly relevant. It is apparent that an advantage lies with countries that from war or post-war experience have the 'know how.' This advantage is exceptionally weighty because of the complete novelty of the engineering problems. It is frequently stated that on account of the difficulty of these problems, costs must be far too great for any commercial development to be envisaged and reference is made to the astronomical sums spent on atomic bomb development during the war. This ignores the fact that in the haste to get results in time to affect the issue of the last war, every practicable method was explored simultaneously and this raised the cost to a fantastic figure. In normal times in developing a new idea for peaceful purposes only the most promising of a variety of possible processes are followed up. Finally, there is the 'wear and tear' or depreciation charge and the rate of interest. The depreciation charge will depend on the success with which the constructional problems are solved, particularly whether the new constructional materials stand up to the radioactive substances involved. The interest charge should be governed by the scarcity of capital in the community. To sum up, Professor Sayers contends that development within the next generation is going to depend on three factors—the success of the search for new materials of

construction, the relative price of coal and other fuels, and the rate of interest. In fact 'a capital hungry world will not be able to afford atomic energy. But a world in which thrift abounds and coal miners are rare birds will jump at it.'

It has been mentioned that the less common metals are assuming increasing importance, but the use of many of them in industry must be severely curtailed by the limited supplies available. Actually it is a vicious circle. New ideas are not developed because of the shortage of the element, and at the same time it does not appear worth while to develop methods for increasing supplies when uses are not apparent. In the search for new metals the vacuum induction furnace has played a major rôle. By this technique the preparation and melting of metals can be carried out free from oxygen, nitrogen or carbon and without contact with ceramic crucibles. The reduction of the iodide of a metal in vacuo, developed experimentally some twenty-five years ago, is being applied on a large scale to yield very pure metals. The reduction of the tetrachloride by liquid magnesium has produced tons of titanium and zirconium, two of the most difficult metals to prepare. The more optimistic predictions for titanium as a high temperature material failed to be realised but, titanium continues to find favour as a metal of exceptional promise in aircraft, marine and armament work. Zirconium is entering a new era. The outstanding property is its excellent corrosion resistance, whilst its very small affinity for neutrons commends its use as a structural material in atomic power plants. The carbides of both titanium and zirconium are also in the news. The first is proving suitable for jet engine applications whilst the second is of interest as a source of zirconium tetrachloride for the production of zirconium metal. Both titanium and zirconium appear to have a future in important alloys which can be made by arc melting in argon or helium atmospheres. Interest is maintained in mixtures of refractory metals and oxides such, for example, as titanium and alumina. It is said that such mixtures withstand higher temperatures than are possible for the metal alone. Amongst the special oxide refractories are zirconia with traces of calcium oxide. This is used in the manufacture of turbo-jet engine blades. Electrolytic methods for the production of titanium are being explored, and there is little doubt that at the accelerated rate at which research is being carried out results of great importance will emerge in this field of the less-known metals. It is significant that in the U.S.A., quite recently, industry got its first chance to see pure hafnium metal. Hafnium is always present as an impurity in zircon sand which is used as a source of zirconium. A development of the utmost importance concerns the rare element germanium, whose existence was predicted by Mendeléeff in 1871 before its discovery in 1886. In 1948 the Bell Telephone Laboratories in the U.S.A. produced a crystal triode—a germanium crystal with two cat's-whiskers spaced a hairbreadth apart on its surface—which when incorporated in a valveless radio receiver operated a loudspeaker. Few could have foreseen that when the early cats's-whisker and crystal wireless set was superseded by the modern sets using thermionic valves that the crystal would again become the focal point of development. With investigations proceeding so rapidly it is difficult to foresee the full implications of this valuable discovery of 1948. However, this new situation calls for supplies of germanium, and hitherto germanium has been difficult to obtain. A British firm has recently perfected a commercial process for the extraction of germanium from flue dust and it is hoped that this country will be independent of imported supplies. The story behind this development is a fascinating one. Some years ago Professor Goldschmidt, a Norwegian scientist, in the course of his geochemical research, examined a piece of British coal spectrographically and found it contained an unusually high percentage of germanium, much of which remained in the ash. Further search revealed that it came from Hartley Colliery, near Newcastle-upon-Tyne, most of whose coal was found to contain recoverable quantities of both gallium and germanium. My old friend, the late Sir Gilbert Morgan, at that time Director of the Chemical Research Laboratory, Teddington, was able to establish that the germanium content

varied according to the source of the coal and that the greatest concentration of this element occurred in the dust deposited in the flues of gas-works producer plants. Upon these observations the new process is founded.

It must not be inferred that the rarer elements are the only elements which are commanding special attention at the present time. Recently at Liége in Belgium nine countries including Britain co-operated in the design, construction and trial of a new type of blast-furnace, which will run on pure oxygen instead of air. Behind this idea lies the possibility of a revolution in iron and steel-making. In practice much will depend on the feasibility of making cheap oxygen in great tonnage. Silicon, which occurs widely in nature is much in the news nowadays. Industry is interested in at least three groups of silicon compounds, the silica sols, silicon esters and silicones. The silicones represent an important range of compounds which since 1943 have been prominent in industrial practice. The production of silicones in Great Britain has hitherto been comparatively small, but supplies will be much greater in the future, and this will promote the wider use of these extremely useful substances. One of their most important properties is their water repellancy. A valuable characteristic of a silicone film is that it is chemically inert and does not deteriorate on prolonged exposure to moisture, sunlight and weathering. These properties are utilised by incorporating silicones in automobile and furniture polishes and in the waterproofing treatment for nylon and other fabrics. The simplest silicones are the methyl silicones, which may exist as oils, resins and rubbers. Resinous silicones have excellent dielectric properties and are used for electrical insulation. They are stable at temperatures at which the usual organic insulating materials decompose rapidly. Silicone rubber may be used for electrical insulation and for other purposes for which rubber is required, especially at elevated temperatures. It will be evident that the chemist has evolved a group of substances with entirely novel properties which are proving most valuable in many ways. Historically it is of great interest to recall that the foundations of organosilicon chemistry were laid by the researches of Professor F. S. Kipping at the University College, Nottingham (now the University of Nottingham). Over the period 1899–1944 he demonstrated the wide range and variety of these compounds. However, the full implications of his work were not generally realised and it was only during the Second World War that the value of organosilicon compounds became fully apparent, largely as an outcome of investigations by American chemists.

Again, there are few known elements which have not been used in attempts to make fluorescent materials. The largest consumer of these fluorescent substances or phosphors is the fluorescent lamp. The world production of these lamps is now at least 100 millions annually, demanding tons of phosphors of the highest purity. Some of the best-known phosphors have their prototypes in natural minerals but there is still much to discover and explain about these remarkable materials. In this field of work the active co-operation of the inorganic chemist, the physical chemist and the physicist continues to extend our knowledge of chemical substances which about twenty-five years ago were laboratory curiosities and are now made on an industrial scale. In recent years, developments have proceeded with great rapidity in the field of fluorine chemistry. The preparation of fluorine by Moissan in 1886 solved one of the major problems in chemistry at that time and between the years 1900–1925 the brilliant Belgian chemist, Swarts, laid the foundation of organic fluorine chemistry. During the past twenty years, workers both at home and abroad have developed methods for the controlled interaction of fluorine and organic compounds. This has given that remarkable category of compounds, the fluorocarbons. A detailed study of CF_2Cl_2 has shown that it is non-toxic, non-corrosive, non-inflammable, has no odour and is not absorbed by foodstuffs, etc. Such compounds—the Freons—are used extensively in the United States in air-conditioning and refrigeration and commercially are the most important group of organic fluorine compounds. Research in this subject is very active and the results

indicate that fluorine chemistry has a bright future.

I have mentioned some of the directions in which developments are expected in the field of inorganic chemistry. Organic chemistry has been characterised by remarkable achievements in the chemistry of natural products and biologically active synthetic compounds and by significant advances made in our understanding of the reactions of organic compounds. Possibly the most romantic chapter in the last fifty years has been that concerned with the unfolding of the story of the vitamins. Probably all the important vitamins have been recognised and most of them prepared synthetically. Indeed several are available commercially. My colleague, Dr. D. H. R. Barton, will later in this session deal with the advancing front of this great branch of chemistry—organic chemistry.

Detailed final results of the 1948 Census of Production are now becoming available and the preliminary examination discloses an illuminating picture of the changing structure of British industry. The census has confirmed great changes in the relative importance of the various industry groups. A shift in productive resources and outputs from industries, such as mining and textiles, towards engineering, vehicles and chemicals is clearly revealed. This is no new phenomenon for it has gone on at least since the beginning of this century and has been an important factor in the industrial progress of Britain. The growth of the chemical industry is spectacular. The industry owes its existence to research and development whilst its growth has been promoted through new markets and the introduction of new products. By starting with a comparatively few substances and combining these into various forms, an ever-increasing range of new products, useful for countless purposes, have been made available. People frequently misinterpret the term 'synthetic.' They consider it means artificial and therefore inferior. Actually, when naturally occurring substances are duplicated by synthesis the justification is higher purities and lower costs. Synthesis very frequently leads to products with unique properties having no natural counterpart. Synthetic fibres, for instance, have made us independent of certain natural fibres and have widened the choice of clothing and fabrics. Synthetic resins are supplementing metals in hundreds of different ways. Synthesis has made possible the thousand and one new chemicals in common use today and will make more available in the future. The chemical industry is, of course, not one industry but literally thousands of specialist industries, differing in the details of the processes and problems of manufacture yet linked by the fact that they are all based on the science of chemistry. It has been a spearhead in the drive for exports and particularly in the groups of drugs, medicinals, dyes, fertilisers and alkalies. This industry faces lively and exciting competition in the world chemical market, but intensive research, manufacture of new products, the use of improved methods as well as a vast programme of reconstruction and expansion, should give it a powerful position. Incidentally, there are many voices raised today to impress on us the importance, the prime importance, of chemistry to industry, agriculture, medicine and defence. It is most desirable that this should be said and said often. There is a real danger, however, that the practical applications of chemistry rather than its theoretical principles and fundamental discoveries should be uppermost in our esteem. Perhaps this is traditional. Alchemy was a very utilitarian pursuit of the elixir of life whereby age and death might be defeated and of the 'philosopher's stone' which would transmute the elements into gold. No one can say what useful practical device may ultimately come from discoveries made purely in the pursuit of knowledge. Applied science draws its life-blood from pure science. Although we may be fascinated by applied science it must never weaken our allegiance to pure science.

I should like now to refer to a special aspect of the advancing front of chemistry. Mr. Winston Churchill put food first, in schemes for a better post-war world. During the past decade, however, most people have had an uneasy feeling that all is not well with their food. As one authority expresses it 'Between the scientist and the bureaucrat eating has lost much of its pleasure and nearly all of its

taste.' A few months ago Dr. Magnus Pyke, who used to be Principal Scientific Officer (Nutrition) in the Ministry of Food, wrote a lucid account of what is done to the food which the townsman eats today. In *Townsman's Food* Dr. Pyke gives some surprising and enlightening information about the chemical processes which are used to preserve or 'improve' food (the inverted commas are Dr. Pyke's own). For example, we are told that it is considered that the increased amount of phytic acid in high extraction flour reduces the human intake of calcium. Therefore, 14 oz. of chalk are added to each 280 lb. sack of flour to balance this loss. Incidentally this action has caused a shortage of the substance used for whitening ceilings. Again we are informed that flour from reasonably clean, new-crop wheat contains up to 18 mouse hairs per lb.; flour from good quality old-crop wheat contains up to 180 per lb. Equally alarming is the widespread adulteration of food. The meat content of galantine and luncheon sausage need be no higher than 30 per cent. 'Fresh fruit standard' jam contains only 20 per cent. of strawberries or 30 per cent. of raspberries. 'Full fruit standard' jam is an artificially dyed product manufactured from a sulphur-dioxide preserved pulp. 'Processed' peas are hard dried peas, taken from store, cooked and dyed green. 'Fresh eggs' may have been in store for nine months. It is salutary that the layman should realise the price he pays for his varied diet.

As we all know, the food processor has chemicals for every need—preservatives, emulsifiers, colouring agents, flavours of all kinds, flour 'improvers,' anti-staling and anti-rancidity agents, sweetening agents, anti-oxidants and antiseptics. Some idea of the magnitude of this problem of chemicals in food can be obtained from the report of the Delaney Committee which was set up in June 1950 by the United States Congress. It was given the task of assessing the dangers to public health arising from the increasing use of chemical substances in food manufacture and from the presence in foods of residues from the pesticides now widely employed in agriculture. The Committee heard the striking evidence from the officers of the Food and Drug Administration, that 704 extraneous substances were thought to be in use. Of these, only 428 were considered harmless as normally employed, leaving 276 of which the safety was open to question. Consideration was given not only to those substances which are added to foods to produce some desired effect but also to those compounds which may become associated with foodstuffs by accident. Examples are pesticides such as D.D.T. which may adhere to crops and find their way into foodstuffs. Experiments were described which demonstrated that D.D.T. appeared in the flesh of chickens, cows and sheep which had eaten crops sprayed with the pesticide. It was present also in eggs and milk. Other pesticides are even more highly toxic. There is now a substantial body of opinion in the United States which favours legislation to deal with the various problems outlined above and this proposal is supported by members of the medical profession and others working in the fields of biochemistry and nutrition.

To appreciate the full significance of this problem of chemicals in food it is necessary to examine the background. During the last fifty years there has been a gradual decrease in home preparation and cooking of food. At one time it was usual for most food to be prepared and cooked in the house. This no longer applies. In fact, nowadays, a great industry has been created which guarantees a steady supply of fruits, vegetables and processed products and is founded on discoveries made in chemistry, biology and engineering. Moreover there is every indication that large-scale preparation, processing and cooking of food will increase throughout the world.

Both in this country and in the United States the legislature has expressed concern at the position. C. A. Adams has pointed out that 'if we cast our thoughts back to the work of the Committee on which our present Preservatives Regulations are based, it will be remembered that the basic philosophy was that chemical substances should not ordinarily be added to food. As a corollary it followed that when chemicals were added their number should be limited to those fully justified by doing a useful job and even then

the quantity used should be controlled.' Perhaps the greatest controversy has arisen over flour. Mellanby showed some five or six years ago that flour with nitrogen trichloride as an 'improver' was toxic to dogs and caused running fits. Rats, guinea-pigs and monkeys, however, were found to be immune and studies in human subjects have also established that man is unaffected by agenised flour. It is wise, however, to be on the safe side and this policy has been adopted by eliminating nitrogen trichloride from flour treatment. Everybody has heard of surface active agents but perhaps their use in food is not always clear. Essentially they are compounds which reduce the surface tension between oil/water interfaces or between a liquid and a solid. Accordingly they find applications in the bakery and confectionery trades, in cakes, buns, ice-cream, salad cream and numerous other products. Some of these substances are used as fat-extenders and it is claimed that by the use of a small proportion of them as much as 50 per cent. of the normal amount of fat required for a food can be saved. Two materials of this type have been the subject of some argument —polyoxyethylene stearate and glyceryl monostearate. Polyoxyethylene stearate has been used in the making of bread and confectionery and whilst it is found to be non-toxic in acute and chronic tests, rather large doses may give rise to diarrhœa. It has been suggested, therefore, the possible effects of the chemical on gastro-intestinal function should be more extensively investigated. The effects of glyceryl monostearate have not as yet been extensively studied and whilst the available evidence suggests it is essentially a non-toxic substance, wider confirmation of this view is very desirable.

There is little doubt that the wide publicity which this subject is receiving will call for planned administrative action in this country. Perhaps it would be of interest to quote from the Fourth Annual Report of the Advisory Council on Scientific Policy. 'The risk to life and health due to the presence of toxic substances in consumer goods is probably small, but the rapid growth of the chemical industry and the needs of the food-processing industries for substances to replace scarce materials or chemical substances which have a claim to use on their own merits as improvers of the appearance, palatability or texture of manufactured foods, are accelerating the pace at which new chemical substances are being introduced into consumer products and the machinery which exists for testing the possible harmful effects of these substances is inadequate.'

From my Address today some may feel that even if chemistry has given with one hand it has taken away with the other. The answer, of course, is not to cancel out the magnificent discoveries I have spoken about but to go forward to still greater achievements which may lead to readjustment. Not all such achievements can be realised in scientific establishments. Some must be the outcome of investigations by the economist, some by the social scientists and some by the legislator. Actually, chemists and chemistry have never enjoyed the whole-hearted approval of people everywhere. As human beings, the chemists are much as other men when distributed statistically over the whole spectrum of human folly and wisdom. Nevertheless, chemistry has made a full contribution to human happiness and as such deserves the support of society. There is quite a considerable and powerful class of people who hold that chemistry is opposed to culture and that a training in the so-called humane studies renders those who have received it broader in mind and sounder in judgment; in fact more fitted for high and lucrative offices. Even schoolmasters have been known to advise gifted pupils that they should avoid science at all costs. As one headmaster expressed it, specialisation in science involved a sacrifice for the public schoolboy for it meant he must give up the study of Greek and in this sense he separated himself from the intellectual aristocracy of his school. Further he considered there was a growing reaction against the teaching of chemistry because it was not sufficiently stimulating intellectually. This prejudice will die but it is difficult to assess the harm it will do in the meantime. The headmaster concerned evades the real and fundamental difficulty which Sir Charles Grant-Robertson states so eloquently in his *British Universities*.

'Every educated man and woman knows well that he or she has been fortunate if in twelve or fifteen years of school and university education they have, among fifty or sixty teachers come under half a dozen men or women who by some unanalysable alchemy of method or inexplicable quality of personality turned routine into inspiration and revealed windows and sunlight in what, but for them, would always be blank and imprisoning walls.'

Some years ago the scholastic agent in Mr. Evelyn Waugh's *Decline and Fall* assured Paul Pennyfeather that it was amazing what one could teach when one tried. 'Why only last term' he said 'we sent a man who had never been in a laboratory in his life as senior science master to one of our leading public schools. He came wanting to do private coaching in music. He's doing very well, I believe.' Today, this satire is too near reality to be comfortable. The lack of competent and enthusiastic teachers of chemistry may well imperil the future of chemistry.

I hope my Address has indicated that the advancing front of chemistry brings with it wide and interesting possibilities of even greater successes in the future. Science can only produce worth-while results if it has worth-while aims, and here I believe chemistry is taking the right direction. In these challenging days when youth stands at the crossways with a confused babel of voices calling for their loyalty and service, I hope the voice of chemistry will not call in vain.[1]

[1] Delivered Sept. 4, 1952.

SOME ASPECTS OF GEOLOGICAL RESEARCH AND THEIR PRACTICAL APPLICATION

Address by

DR. J. E. RICHEY, M.C., F.R.S.

PRESIDENT OF SECTION C

IN the English-speaking world alone very many geologists are at work at the present time in the field and laboratory exploring the mineral constitution of the earth. Many are engaged in mapping new territories and, in general, wherever the place may be, are carrying out fundamental research on stratigraphical, structural, and other problems of varying complexity. Many are concerned with economic enquiries, applying the normal methods of geological investigation to the elucidation of practical issues, and aided by new techniques which have been developed for specific purposes. In addition to pursuing their own researches, many are employed as teachers and lecturers in schools and universities. I do not propose to touch further on the educational aspects, important though they are in providing an outlook on the world around us wider and fuller than perhaps any other science can give. Later in my address I shall be dealing briefly with some problems on applied research, a field in which geology is playing an increasingly important part. In this connection I should like to remind you of the varied content of the phrase 'economic geology.' We have only to think of its many divisions and subdivisions, its numerous applications in the coalfields, in the search for oil, in metalliferous mining, in agriculture, forestry, water-supply, and in civil engineering undertakings with all their multifarious ramifications, to realise how important a share geology is taking, directly and indirectly, in the control and development of world affairs. So varied, indeed, are the branches of applied geology and so many the contacts with other sciences, as, for example, in geophysical research, in the development of natural resources, using that term in its broadest sense, in land utilisation problems, that the geologist nowadays is often compelled to become a specialist in a particular field. This specialisation, indeed in some cases ultra-specialisation, brings with it its own dangers, in that the geologist concerned may lose contact with the main currents of geological thought and may find himself in a by-water leading to a dead end. In both pure and applied geology too narrow a view and too restricted a field of observation have led in the past to premature and wrong conclusions. Some examples of this are given later but one additional instance may be noted here. We may recall the early views concerning our 'Northern Drifts' and how Agassiz with his wide experience of the action of land ice placed the interpretation of glacial deposits in the proper geological perspective.

There is just one further word I should like to say in regard to pure research by way of introduction. As Henri Poincaré remarked in *La Science et l'hypothèse*: Science is built of facts, as a house is built of stones; but an accumulation of facts is no more a science than a heap of stones is a house. My object in this address is, largely, to indicate manners of building the 'geological house' which have been followed in certain well-known investigations. For, as expressed by W. H. George in *The Scientist in Action*: 'The art of making discoveries should be extended by considering noteworthy examples of it.'

GEOLOGICAL METHODS AND PRINCIPLES

The geologist is concerned fundamentally with observation, that is, with the

systematic recording of phenomena as they present themselves to him. He has also to bring together his observations into a pattern of evidence, from which a logical conclusion may be drawn; or, to quote from Professor A. N. Whitehead's *Adventures of Ideas*, he has to 'search for simple statements which in their joint effect will express everything of interest concerning the observed recurrences.' There are, then, three stages in a piece of research which should be clearly stated and kept in view: (1) observation or the gathering of facts; (2) the framing of a conclusion to explain the facts; and (3) synthesis, i.e. the application of a series of conclusions so as to establish a broad general explanation or to enunciate a principle. I should like to make it clear that in what follows I am considering only what I call the patterns of evidence which lead the geologist to his results, not the technical aids he brings to his assistance. Nor do I propose to lay down rules of evidence. The method of proof, that is of reaching a conclusion upon a body of data, is varied according to the circumstances, and each geologist uses his own judgment to a great extent in the matter. In my experience, and probably in that of others, the various kinds of proof-patterns are utilised almost subconsciously and are often intertwined. As illustrations, I propose to refer to a few notable investigations carried out at various times during the past century and a half.

Hutton's fundamental researches in the latter part of the eighteenth century are altogether admirable in their ways of approach, however difficult his manner of writing may have been. He brought to bear upon his problems a wealth of facts to which he always refers his immediate conclusions, as being the arbiters in the case. In an age when the Catastrophists appeared supreme, it is all the more instructive to find how complete their overthrow by the Uniformitarians eventually was, in face of Hutton's facts. His explanations were based mainly upon natural processes which he could observe in action in the world around him. His only further requirement was time for the results recognised in the rocks to be produced, though time of such a length that his realisation of it was one of the greatest strides into the unknown which anyone has ever made. I am not setting out to review the philosophic aspects of Hutton's work, or the full scope of his *Theory of the Earth*, with its recurring cycle of earth change. But I wish to consider very briefly his methods of approach to three fundamental problems: the significance of unconformities, the magmatic origin of basalt, and the consolidation of granite from a molten state.

Already, in 1667, Nicolas Steno, the Dane who settled in Florence, and in 1719 John Strachey in England, had recorded their observations of the arrangement of strata known as unconformity, but without explanation. Hutton, however, rediscovered and interpreted the relationship, and it is the manner of this advance which I wish to emphasise.

In the south-east of Scotland where Hutton mainly worked, he was constantly confronted with rocks of two kinds. Both, he realised, had been laid down as flat-lying beds of sediment upon a sea-floor or other area of deposition. Further, two contrasted attitudes of the bedding were apparent, each related to rocks of different composition, one of them highly inclined, the other in flat or gently-inclined layers. The question that arose in Hutton's mind was: what would be seen at places where these two contrasted rocks came in contact with one another? His search for such critical exposures was rewarded by the discovery of the unconformity along the River Jed near Jedburgh and at Siccar Point on the Berwickshire coast. At these places the flat-lying beds rested upon the edges of the highly-inclined strata. So much was observed. The solution became equally clear. The junction of the two series of rocks represented a former surface of the earth. Sediments laid down in the sea had become first compacted and later uplifted by earth movement. Next, they had been shorn across by a slow process of subaerial erosion such as Hutton could see in action around him. Lastly, the final surface of denudation of the ancient land subsided in turn beneath the waters of a sediment-forming area, from which the upper, unconformable series of beds were deposited, and so a new cycle of earth change

was initiated. What was the essential pattern of this research? I would say that the realisation of two contrasted sets of circumstances, of vital differences, was the main pattern leading to the discovery, and secondly, the appreciation of the adequacy of natural processes at work at the present day to explain an observed situation.

Much had been achieved by such men as Guettard and Desmarest working at different times in the Auvergne to show that basalt was of magmatic origin before Hutton and his associate in Edinburgh, Sir James Hall, published their results. Hall in particular established a new approach. He recognised in 1785 that basalt filling highly-inclined fissures in the ancient crater wall of Somma bore fine-grained margins, which he later explained to be the result of chilling of injected lava against the cold sides of the fissure. Early in the nineteenth century similar fine-grained and compact marginal zones were observed by Thomas Allan, in association with Hall, along the base and top of the great basic sill forming Salisbury Craigs, Edinburgh, and on the sides of a basalt dyke at the Cat's Nick traversing the sill, and their correspondence with those of Somma was realised. The phenomenon of the intrusion of basalt in areas where no active volcanoes were in existence was established by these observations. Hall had already, in 1798, adduced evidence from the laboratory that basalt was of magmatic origin, by melting basalt and subjecting the melt to slow cooling, so partly reconstituting the crystalline state of the original rock. In these examples the method of proof lay in making comparisons and recognising similarities, in other words, by analogy, between what could be demonstrated by experiment or in nature and the matter at issue.

It is appropriate to recall the opposition which Hutton and Hall's conclusions in regard to the magmatic origin of basalt had to meet. Rocks cited as evidence against the igneous theory will be examined by members of the geological section during the present Meeting of the British Association. Kirwan, president of the Royal Irish Academy, a fervent supporter of Werner's Neptunian school, pointed to the occurrence of basaltic rocks at Portrush which contained fossils and therefore were of aqueous origin. These rocks, however, though superficially resembling basalt, are highly-indurated Lias shales in contact with a dolerite sill which has altered them.

I may refer briefly to another great contention of the Hutton school, that concerning the origin of granite. To Hutton, the discovery of veins of granite crossing the rocks of Glen Tilt in Perthshire indicated that the granite had been molten and so capable of being injected into fissures where it solidified. His reasoning was sound and no one has convincingly disputed it. It led him, by the rule of similarity, to envisage large masses of granite as also having been molten. Most geologists agree that granitic rocks may also be formed by the granitisation of country rocks with or without the visible intervention of a magmatic stage. Professor H. H. Read has recently emphasised in an Alexander du Toit memorial lecture that granitic studies must be related to their regional environment. Thus, the development of large granitised bodies, he maintains, can only take place in regions of high metamorphism. In low grade metamorphic areas granitisation will be localised and marginal to intrusive granitic bodies.

The Law of Superposition of strata provides evidence of their relative ages, and was already established when William Smith, the English land surveyor, recognised that the varied Jurassic rocks of the west of England succeed one another in a regular order. Had he not been an assiduous collector of fossils as well as a stratigrapher his great discovery of order in fossil-contents as well as in lithology could not have followed. By the end of 1795 he had become convinced of the similarity of fossils obtained from the same set of beds in different places, and also realised that successive groups of strata had provided different fossil-assemblages. His knowledge of the time-order of the strata enabled him, from his realisation of the patterns of similarity and dissimilarity among the fossils, to make the greatest of all stratigraphical discoveries, that fossils are a sure guide to correlation of strata over wide areas. As he expressed it in a

letter dated January 5, 1796,—' that most wonderful order and regularity with which Nature has disposed of these singular productions and assigned to each Class its peculiar Stratum.' (See L. R. Cox: *Proc. Yorkshire Geol. Soc.*, 25, i, 1942.) His discovery has added point to the truism that life is the one non-uniformitarian element in the world's history, ever changing and never exactly repeated.

Lapworth's genius for appreciating the kernel of a problem was strikingly manifested in two areas. At Dobb's Linn near Moffat, in the Southern Uplands of Scotland, his discovery that a series of graptolite assemblages was repeated in inverted order in what appeared to be a straight succession of shales, established the stratigraphical succession and simultaneously determined the typical structural unit of the region. In the N.W. Highlands, by mapping and by subdividing previously known lithological groups of strata, he was able to show that an apparent stratigraphical succession was fallacious, and that other criteria were required for resolving the problem. His work emphasised the need for a stratigraphical succession to be clearly demonstrated as a pre-requisite for the elucidation of structure, and specifically, that an apparent succession of strata may be misleading and the Law of Superposition inoperative in certain terrains.

In the two cases, the pattern of discovery was somewhat different. In the Southern Uplands Lapworth recognised that the two-way graptolite succession did not fit in with any previous explanation, and proved that the repetition of the succession was the result of folding. In the N.W. Highlands he realised the quandary into which Murchison's ' facts ' based upon the Law of Superposition led. To quote his own words, ' they land us in a set of conclusions so unnatural and absurd that it would be ridiculous to suppose that any scientific man could entertain them for a moment.' The *reductio ad absurdum* method of criticism has had its applications in geology.

The ' unnatural and absurd ' conclusions were that the Cambrian-Ordovician series, to quote Lapworth's summing up, ' is composed of two members [at Durness] and of three members [at Eriboll] and the well-known Durness Limestone forms both the central and the highest of these. This strangely constituted fossiliferous series is both older and younger than the Sutherland gneiss, for it underlies the latter conformably and overlies it unconformably.'

Lapworth invoked folding of alpine type, as enunciated by Heim, and large-scale lateral displacements of rocks to account for the anomalies in the apparent succession. Although his detailed examination of a limited area did much to settle the N.W. Highland problem, one cannot read the literature of those times without paying tribute also to two other geologists. One, Professor Nicol of Aberdeen, in 1861, had already advanced very far towards a solution, for example, by finding evidence of complete inversion of the upper of Murchison's two quartzites. The other, Dr. Callaway, accurately depicted in his sections published in 1883 (the year of Lapworth's paper ' The Secret of the Highlands ') great displacements of the rocks along low-angled planes. For example, he depicted the two spectacular thrusts along Loch Glencoul, the lower displacing Lewisian Gneiss over Cambrian, the upper bringing forward the Eastern Schists along the Moine Thrust, as it was later termed.

The subject of igneous tectonics covers a wide field, and before discussing particular Hebridean features I feel that some tribute should be paid elsewhere. I may mention more particularly G. K. Gilbert's early recognition of laccoliths in the Henry Mountains east of the Rockies, Hans Cloos' many field techniques developed in the Reisengebirge and other areas for elucidating the structure of plutonic and volcanic masses, and Professor Reginald Daly's stimulating and comprehensive work throughout a long life. In igneous as in sedimentary tectonics a geometrical pattern of evidence is often apparent. In the case of the Tertiary intrusive complexes of the Inner Hebrides, Harker working in Skye presented in 1900 the first pattern to be based on detailed mapping. An apparent simplicity of structure of the plutonic masses was attributed to laccolithic intrusion,

incorrectly as we now know. It is in a way paradoxical that the almost unbelievable complexity of Mull should have yielded the main geometrical patterns applicable to most of the Hebridean-Irish region. In one important aspect the research was furthered by the prior discovery, in 1909, of the cauldron subsidence of Glen Coe. There a cylinder of Lower Old Red Sandstone volcanic rocks, six miles in diameter, bounded by a ring-fault, was let down among Highland schists for some thousands of feet. Along the fault two ring-intrusions of granitic magma found their way upwards from an underground source. It was fortunate that Clough and Bailey, who with Maufe mapped and resolved Glen Coe, were foremost in the investigation of Mull. Tribute is also due to E. M. Anderson for his interpretation of the origin of the contrasted crustal fractures governing the geometrical patterns which mapping disclosed. The part which crustal fracturing plays in igneous intrusion, yielding curving rock-outcrops arranged around successively functioning centres, was fully demonstrated in that complicated district. In brief, ring fractures were of two kinds. One of them, related to intrusions mainly of plutonic type, was similar to that of Glen Coe, being vertical or outwardly inclined from the centre (ring-dyke); the others were inclined inwards towards the intrusion-centre and were exemplified by inclined sheets, or cone-sheets, mainly with hypabyssal crystallisations. The demonstration of curving intrusions, whether ring-dyke or cone-sheet, related to geometrical centres, and the migration of a centre from one point to another during the building up of a single complex proved fruitful of results in other parts of the Hebridean-Irish region, at Ardnamurchan, in the Mourne Mountains, at Slieve Gullion and at Carlingford. At Slieve Gullion a ring-shaped intrusion of dyke-like form had been mapped by Joseph Nolan and other Irish surveyors some years prior to Harker's arrival in Skye, although it was not separated till recently into its component intrusions or from others of related rock-type and sheet-like form which enter into that interesting complex. Recent discoveries of the Hebridean ring-dyke pattern have been made in other countries and the idea of relationship of fracturing to centres concomitant with intrusion helped much in their elucidation. Especially noteworthy are the magnificent series of ring-dykes mapped in difficult terrain in New Hampshire in the north-east of U.S.A. by Miss Louise Kingsley, Professor Marland Billings, the two Chapmans and David Modell. Cauldron subsidences in the Oslo district mapped by Schetelig prior to the survey of Mull have recently been described by Oftedahl, Professor Holtedahl and Saether. The precision with which the amounts of subsidence of the central pre-existing rocks have been determined for some of the Oslo complexes is one of the most impressive results attained in that classic area.

Cone-sheets, inwardly inclined towards an underground focus situated at a depth of several miles below the present surface, are found in varying abundance at every one of seven or eight Hebridean-Irish complexes, but as far as I am aware nowhere else in the world in association with ring-dykes. Yet in the Inner Hebrides especially, the two contrasted types of intrusion appear to have been complementary to one another, their intrusion alternating repeatedly in time in Mull, Ardnamurchan and elsewhere.

The geometrical pattern of the Tertiary types governs conclusions in regard to methods of intrusion, as it sheds light on the shapes of the spaces occupied by the rocks. The problem of how spaces originated for intrusion is mainly in the first place one of geometry, for without some precise idea of the shape of the space concerned there can be no sound basis for argument about how the space originated. The difficulty of the subject, even in the case of minor intrusions, is apparent to anyone who has tried to visualise on some geometrical pattern the shapes of repeatedly cross-cutting quartz-dolerite sills of the Scottish coalfields.

Great discoveries in geology which are dependent upon inductive reasoning are often made from the observation of several phenomena all of which point in the one direction. The first example I advance is the explanation proposed by Wager and Deer in regard to the Skaergaard Gabbro in eastern Greenland, that the mass

consolidated in the main from the bottom upwards and that the process was governed by convection currents set up in the liquid rock owing to top cooling. It is unnecessary, I am sure, to detail the many lines of evidence which led the authors to their classic explanation: structural, petrographical, chemical and mineralogical features contributed to the solution, and all pointed in the one direction.

The second example is Professor Kennedy's conclusion that the Great Glen Fault which divides the Scottish Highlands in two along the line of the Caledonian Canal and Loch Linnhe is mainly of transcurrent or wrench type, and that the rocks on one side of the fracture have been shifted laterally for 65 miles relatively to those on the other. Three major investigations preceded the enunciation of the conclusion, and with at least one other main line of evidence contributed to the solution. Further work may reveal other lines of evidence but, whatever they are, they will have to take account of a series of major coincidences which, I think, add up practically to a demonstration of the most powerful wrench fault known.

Patterns of Evidence

I am aware that the patterns of evidence I have adduced for the researches selected are of different values and types, and that investigators make use of a variety of patterns in a single piece of research. Nevertheless I am venturing to mention and briefly consider these and other patterns as a conclusion to what is necessarily a very incomplete enquiry.

Anomalies.—A pattern of evidence may include an anomaly at variance with apparently valid conclusions reached on other grounds. Anomalies may not constitute vital objections to a conclusion: on further investigation they may yield and conform to the prior solution, or prove irrelevant. On the other hand the recognition of anomalies has frequently been a most fruitful source of discovery in geology as in other sciences. It is not so easy as it may appear to recognise an anomaly of critical importance during a piece of field research. A difficulty encountered may be disregarded when it is at variance with a mass of data pointing to a particular conclusion. Yet the recognition and tackling of such a difficulty often lead to a new and entirely unforeseeable point of view.

An anomaly which helped greatly towards the elucidation of Dalradian tectonics in the Scottish Highlands is that of the three-sided stratigraphical formation. Maufe in the first instance recognised the problem presented by the Ballachulish Limestone, which along different parts of its outcrop was in contact with three other Dalradian formations. Together with other evidence this led Sir Edward Bailey to his well-known interpretation, that the rocks were involved in large-scale recumbent folding accompanied by fold-sliding, so accounting for the anomaly of the three-sided formations in the Ballachulish district and elsewhere in the Dalradian Highlands.

Similarity and Dissimilarity. The visual recognition of the like and unlike is part of the geologist's all-day and everyday experience. In larger issues, leading to some new discovery, it also plays a part. The noting of a change of appearance or mode of weathering, however slight, in a traverse across a granite takes the investigator back on his steps in order to ascertain a reason for the change. A sharp, intrusive contact of one type against the other may be found, as in the case of the Mourne granites in Co. Down where four or five different intrusions were recognised largely from marginal differences in aspect of the rocks. In a still larger mass, the Cairngorm granite north-west of Deeside, no change of type has been observed, and argument from similarity requires this to be regarded as a single intrusive body, by which is meant a homogeneous mass in which consolidation went on without a recognisable break.

Analogy.—Comparison with a well-founded case from another locality is a familiar means of arriving at a solution or is often cited as supporting evidence. There are two requisites: the standard case should be well documented and beyond dispute, and the comparison should be relevant, i.e. the two cases should be alike. Analogy is most effective either if the standard case is personally known to the investigator or if the

phenomena on which the standard case is based are such that the investigator can assess their value from his knowledge of similar circumstances. Much care and discrimination are needed in applying the method, and the use of analogy to strengthen a position often requires a singularly unbiased mind. Major discoveries should be based adequately upon the local evidence, but in minor ones with strictly limited conclusions the method of proof by analogy is often justified.

Repetitive Evidence.—The observation of repetition is perhaps the most usual means to an end in geology. A single case of a phenomenon may be interesting and intriguing, but unless many instances are found there may perhaps be no adequate pattern of evidence to suggest a problem or to provide a solution. To cite a Scottish example : in 1902, in the Craignish peninsula south of Oban, Maufe found dykes of two different compositions, each set following different directions. One, composed of basaltic rock of Tertiary type, extended north-westwards in the direction of the Island of Mull. The other, composed of porphyrite and of other Lower Old Red Sandstone types, was directed north-eastwards towards the granitic complex of Loch Etive. Maufe also noted instances of the intersection of north-east dykes by those of the north-west set. At right angles to their directions of trend the dykes of both series become less closely spaced and finally are absent. Now, a dozen dykes would have raised no particular question, but the delimiting of the two swarms directed respectively towards Mull and Loch Etive, with dyke-free country on either side, carried with it the suggestion that the intrusion of the dykes was related in some way to these two intrusive complexes. This type of coincidence led to the explanation given by Bailey, that dyke-fissures became concentrated across a weak place in the earth's crust where, as indicated by plutonic intrusions, large bodies of magma were locally in existence below ground.

Geometrical Pattern.—The recognition of a ruling geometrical pattern is a prevalent aid to geological mapping. Its discovery in almost all cases gives greater confidence in drawing outcrops on the field map. The pattern is really an expression of the geological structure. For example, the discovery of an isoclinal fold at Dobb's Linn in the Southern Uplands by Lapworth provided the type of pattern for mapping the greater part of the region. Reference may also be directed to fault patterns, where prevalent fault directions are recognised as characterising a particular district. In coalfield surveys where the working out of the structure so often has to depend upon information supplied by bores, with few or no surface outcrops, it becomes a matter of special importance to know what the prevalent direction or directions of faulting may be, so that the information can be examined in the light of these trends.

Intersecting Lines of Evidence.—In most of the immediate problems which constantly concern the geologist a single line of evidence may prove adequate for arriving at a conclusion. But it may well be, especially in problems of magnitude, that at least three converging lines are necessary to reach a final conclusion. No general rule in the matter, however, can be given, and every worker must decide for himself the needs of his problem.

Practical Applications of Geology

The practical importance of a knowledge of the geology of a country has long been recognised. For example, in the British Isles, as early as 1813, Macculloch was appointed to carry out a geological survey in Scotland in connection with the work of the trigonometrical survey. In 1832 a similar appointment was made in Ireland in the person of Captain (later General) Portlock, R.E., who, in 1843, presented his notable Report on the Geology of the County of Londonderry and parts of Tyrone and Fermanagh. In the preface Portlock states that the object in view was to combine with scientific investigation much practical information. This dual purpose, widened to include investigations into mineral resources, has been the guiding principle in shaping the policies of national geological surveys at home and abroad, although on occasions the need for carrying out special economic enquiries has taken precedence over comprehensive mapping. It is hardly necessary to say that a thorough training in

geological principles and in their application to particular problems is an essential pre-requisite for success in the economic field. A further point which should be specially emphasised is that economic or applied geology is not a separate subject from geology. It is simply geology applied to a variety of objectives containing an economic purpose. In amplifying this basic truth by quoting certain cases, I shall attempt to indicate ruling patterns of the evidence.

Coalfield Geology.—Amongst the wealth of scientific achievement which has come from the study of coal it is difficult to single out a subject in illustration of the main patterns of evidence involved. But perhaps the discovery and exploitation of the concealed coalfield of Kent in the south-east of England may be selected, for a variety of reasons. In the words of the late Professor Boyd Dawkins, spoken in 1906, its discovery is 'the story of a scientific idea, originated many years ago, taking root in the minds of geologists, developed into theory, and ultimately verified by facts.' The pattern of evidence was largely one of geometry, though geometry on the grand scale. Even in day-to-day coalfield research, the same type of evidence is constantly required in dealing with structural problems, and the conclusions reached serve as a guide to the location of bores.

In 1821 Buckland and Conybeare drew attention to the similarity of the structure of the coalfields of Bristol and Somerset on the one hand and northern France and Belgium on the other. In 1846 De la Bêche stated that there might be Coal Measures beneath the intervening Jurassic rocks of the south of England, and in the same year a bore for water in the Pas de Calais district proved Coal Measures on the other side of the Channel. In 1856 Godwin-Austen gave more precision to these early ideas. He visualised a far-extending belt of uplift that had involved the Coal Measures and older rocks, stretching from the south-west of England to northern France and Belgium. Southwards against the uplifted belt, after the removal of the tops of anticlines by post-Coal Measures denudation, he suggested that the cover of younger strata would have overlapped and at the same time become thinner than farther north. So the cover, overlying Coal Measures preserved in troughs, might be sufficiently thin in the south-east of England for the underlying coals to be reached by boring and exploited.

The idea became accepted by many though not all geologists, and the experimental stage of putting the matter to the test by boring was begun in 1872 near Battle, north of Hastings in Sussex. The bore, however, showed that in this area the upper oolites of the Jurassic by themselves were much thicker than anticipated, and the bore did not extend below the Oxford Clay. But the idea persisted, though as now known it was incorrect as regards the direction in which the thinning of the Mesozoic cover takes place, and in the mind of Boyd Dawkins, in collaboration with the engineer, Brady, led to the drilling of a bore starting in the Chalk close to the Shakespeare Cliff near Dover in 1886–90. The bore reached Coal Measures at a depth of only 1,100 ft. and proved a number of seams of coal. Many more deep bores were subsequently drilled in Kent, and exploitation at several collieries followed the experimental stage. As H. G. Dines and others have recently shown, Coal Measures are now known to occur within an area of about 100 square miles bounded by Dover, Canterbury and Ramsgate in east Kent, in a structural basin the axis of which extends west-north-west through a point a few miles north of Dover.

Exploitation has been accompanied by much geological research, and classification by means of widely applicable methods of comparison has proved useful in Kent as in so many other coalfields in Britain and elsewhere. In this study H. Bolton working with the animal remains and Newall Arber with the plants applied stratigraphical knowledge gained elsewhere to determine the position of the Kentish Coal Measures in the general sequence. Later work by Sir Arthur Trueman on the non-marine lamellibranch faunas and by Stubblefield and others on the marine horizons has modified the original classification of the lower part of the succession. The two seams of coal mainly worked at present are situated in

the Upper Coal Measures, the lower seam towards their base.

The less spectacular though on occasions highly profitable method of detailed comparison of lithological sequences should also find a brief mention here. The life of a coalfield depends largely upon the number and lateral extent of its workable coal seams. The correct recognition of the various coals in a local succession thus becomes a matter of vital importance in ground so much broken by faults as many of the British and Continental coalfields are. It would be invidious perhaps to select as an illustration one of the many cases where Survey geologists and mining engineers have corrected a wrong designation of a seam and have brought welcome extensions of life to particular coalfields. The value of their work emphasises the need for a continuing study while coal remains an essential asset for Britain's economic prosperity, and here I may refer to a recent Presidential Address to Section C by Dr. Murray Macgregor (Dundee 1947) in which many other desiderata for coalfield research have been indicated.

The Search for Oil.—Much has happened since the year 1859 when the first oil-reservoir to be discovered was pierced in Pennsylvania by ' Colonel ' Drake's drill at a depth of 69 feet. Scientific methods of locating possible reservoirs have been developed with such precision that nowadays deeper and deeper sources of oil are being sought and exploited. In the United States alone, in 1945, no fewer than 65 wells were drilled to depths greater than 12,000 feet. In the early search for oil geology was hardly used at all, but the phenomenal developments of the past 50 years have resulted very largely from geological and geophysical investigations. At the present time in the United States the chances of locating an oil-reservoir by the ' wild cat ' drilling that was almost exclusively practised up to the close of last century have been estimated as only 5 in 100, as compared with the chances of 23 in 100 for cases in which geological and geophysical methods have been employed.

The geologist and geophysicist are concerned mainly with the determination of structure, and geometrical patterns of evidence are of chief value in the discovery of suitable ' traps ' for oil. The anticlinal theory of the early days became well substantiated with the inception of organised field research by geologists, and geological mapping increased in importance from 1897 until 1923 when, in the United States, the first oil-location based upon a geophysical (gravity) survey was drilled. To illustrate an idea which has proved highly profitable, I propose to take an example not from the U.S.A. but from the other world-pole of oil production situated around the Persian Gulf.

Early Tertiary (Eocene) beds form the surface rocks of Bahrein Island, 30 miles long by some 10 miles in width, in the south-eastern part of the Persian Gulf. The rocks were examined in 1908 by G. E. Pilgrim of the Geological Survey of India who showed that the strata occurred in the form of a slightly elongate dome. He noted associated deposits of asphalt and speculated on the probability of oil occurring at depth. His idea was not put to the test till much later. The delay no doubt is attributable to the fact that the Bahrein rocks lie at a lower stratigraphical level than the oil-producing beds—the Asmari Limestone—of the adjoining great Persian field. It was not until 1932 that the Standard Oil Company of California discovered by drilling the first oil-pool at Bahrein in Cretaceous limestone. Since then production has gone on apace, and reserves in this small insular area have been estimated to be nearly one-half of 1 per cent. of the total known world resources of petroleum. An interesting feature of the production side is that the pressure in the oil below ground is maintained by injecting into the oil-reservoirs high-pressure gas obtained from a level in the Cretaceous deeper than the oil-horizons concerned.

This is not the place to comment upon the variety of structural and other traps which the world-wide search for oil has disclosed. But it is perhaps of interest to refer to another aspect, namely, the importance of voids in reservoir rocks, and the chances which have proved successful of finding oil in dense strata owing to the local occurrence of fissures and fractures. The recent discovery of oil in northern Venezuela in dense Cretaceous limestone

below previously exploited porous Tertiary strata provides an example of the rôle of fracturing associated with folding. Another case is the Balkassar field in the Punjab. The search for fissured zones at depth should prove to be a fruitful if difficult subject for future investigations, and in their prosecution geological and geophysical research may indeed pay high dividends.

Appreciation of the conditions under which the initial migration of oil takes place from the source rocks to the porous or fissured reservoir rocks is also vital. The matter was discussed at the Geological Congress in London in 1948 by L. G. Weekes, and emphasis was laid upon the importance of unconformity and overlap, in other words, the positions of ancient shore-lines with off-shore dip of beds, for these are guides to the present-day geographical locations where oil may be expected to occur. In resolving such problems of palæogeography every available branch of geological science must be brought into action, and the future holds no doubt as high a place for geology in the search for oil as the history of the past fifty years testifies.

Location of Ore Deposits. Exposed ore-bodies yield denuded fragments of ore to adjacent streams and rivers and were discovered by old-time prospectors finding these clues along the waterways and following them up to the parent masses. Less obvious geochemical and biochemical methods are now being explored for tracking to their sources the superficially oxidised and water-soluble derivatives of hidden primary ore-bodies. These methods include the field analysis of surface water, and laboratory analysis of parts of plants and trees and their derived humus, since heavy mineral salts in soils can be concentrated in the tissues of certain species. An ore-body suspected for one reason or another in a certain place may be located by geophysical or other means. Geology becomes highly important, however, in hunting for fresh ore in fields which are under exploitation. These provide much more valuable geological information about the mode of occurrence of ore than natural exposures alone can do. A plea may be made here for the retention of geological advisers throughout the working of a mine. By this means as the workings progress valuable information can be gathered which may be vital for the understanding of the ore occurrence but may be destroyed during exploitation.

The literature of ore deposits contains numerous examples of the successful application of field geology in the search for ore-bodies. The hydrothermal solutions from which many ores have been deposited often followed structurally controlled channels. Generally speaking there are two major types of structural control, namely, folding and faulting. In particular, the recognition of the geometrical pattern of fractures has proved to be of first-rate importance in locating new bodies of ore or in extending the knowledge of veins already found. The simplest fracture-pattern is the parallel or sub-parallel system, as seen at Cœur d'Alene, Idaho. More common, however, are combinations of different sub-parallel sets. These may intersect to form a roughly rectangular pattern, as at Real del Monte, Mexico, or a branching pattern, as at Oruro, Bolivia. Less frequently, more than two sets of veins are present. At Butte, Montana, this is the case, and although the vein network is complicated an underlying regularity can be made out.

An instructive case where modern methods of mineralogy and petrology have been used in the search for ore comes from the important mining locality of East Tintic in Utah, and is concerned with the alteration of country rocks. Here rhyolitic lavas, probably of Oligocene age, rest upon highly-folded Palæozoic carbonate rocks and quartzite, and both rock series are cut by monzonitic intrusives. Ore occurs only in the rocks below the lavas, and in consequence may be ' blind.' About 1927 Billingsley made recommendations based partly upon the strong alteration of the barren lavas, and these led to the discovery of the North Lily ore-body. This early work was not followed up until the last war, when the United States Geological Survey studied in detail the time and space relationships between rock-alteration and ore-deposition in the area. From these studies Lovering and his collaborators established

five stages of hydrothermal alteration : (1) early barren stage, (2) mid-barren stage, (3) late barren stage, (4) early productive stage, and (5) ore stage. The first stage is a dolomitisation of limestone and chloritisation of lava. Mapping of dolomitised fractures in limestone gave valuable information about the hydrothermal ' plumbing ' system, for although the early barren stage solutions travelled far beyond the ore solutions the latter generally followed channels previously used by the earlier. The mid-barren stage is one of argillation and is not clearly related to ore deposition, but the succeeding late barren stage, characterised by jasperoid and pyrite, is intimately associated with ore, and the miners consider 'quartz ' to be an excellent guide to ore. Sericite and hydromica of the early productive stage is a most useful guide, but the sericite must be distinguished from that of the mid-barren stage. Although Lovering states that ' no alteration thus far recognised is a sure indication of ore,' yet his work has borne important fruits and opens up fresh lines of thought in ore geology.

Civil Engineering Problems.—In illustration of the methods of approach in geology to civil engineering problems, I propose to describe two examples from Britain. One concerns the foundation of a dam, and is a further instance of the successful pursuit of a scientific idea. The other, the site of a tunnel, was solved by a pattern of converging lines of evidence.

The investigation of the Vrynwy damsite in North Wales in the '80's of last century is an example of the use an engineer made of his knowledge of geological processes. One of the engineers concerned in the scheme, Dr. G. F. Deacon, was aware that the valley had been occupied by a glacier, and that erosion by ice produces unevenness in rock-floors. He recognised, too, that the superficial deposits covering the valley-bottom had been deposited in a lake of glacial times, and surmised that the lake had been dammed by a rock-bar. To test his idea, nearly 200 bores in all were put down through the surface ' soils,' and a rock-bar was eventually located at a depth of 40 to 45 feet. Rockhead sloped away on either side of the bar, which was used for the foundation of the great masonry dam, and Deacon estimated that, had a site been chosen at a distance of only 220 yards away, this would have involved an additional cost of no less than £300,000 to £400,000 for the works. In similar circumstances at the present day a geophysical survey should eliminate the need for so extensive a boring programme, but we may admire Deacon's persistence in his search and his well-deserved success.

The second example is the investigation of the site of the Mersey Road Tunnel, known as the Queensway Vehicular Tunnel, by Professor P. G. H. Boswell, which was opened to traffic in 1934. It is a striking instance of the reliance placed upon geological inference which could not be verified by trial boring owing to the rapid ebb and flow of the tide. From the driving of the Mersey Railway Tunnel begun in 1879, and from investigations prior to that undertaking by the engineer-geologist, T. Mellard Reade, it was already known that the River Mersey was flowing in a buried river-valley or rock-channel which was lined with a deposit of boulder clay. The new tunnel was sited a short distance downstream (i.e. seawards) of the Railway Tunnel. This raised the question whether the rock-channel was deepening seawards, as might be expected on general principles, or landwards. The investigation took account of the knowledge that the Irish Sea was filled with ice when the rock-channel was cut, and the exit of the Mersey seawards in consequence blocked. It was also known from bores in the Liverpool and Birkenhead area that tributary buried channels existed and that, in plan, their courses made V-shaped angles with the main channel pointing inland. As with all normal tributaries, this therefore had been the direction of river-flow. Further, the depth of the main buried channel of the Mersey was known from bores farther east to be considerably greater there than the depth proved during the excavation of the Railway Tunnel. From these three main lines of evidence Boswell concluded that the river responsible for the erosion of the rock-channel had flowed landwards, not seawards, and that its floor at the site of the Road Tunnel would be at any rate not

lower than at the Railway Tunnel. The deduction turned out to be strikingly correct, and so near to rock-head was the Road Tunnel driven that at one point in the excavation only 3 feet of rock separated the tunnel-roof from the bottom of the rock-channel.

Converging lines of evidence are perhaps the most usual means of solving a major geological problem in civil engineering, although approach by analogy, from his experience of the surrounding or similar terrain, may often be most useful to the geologist. As the problems that arise are as diverse as geology itself, and the situation different to a greater or less extent in every case, experience is undoubtedly a main asset for success.

In conclusion, I may perhaps venture to add a few words concerning the investigator himself. It is common knowledge that discovery waits upon the coincidence of time, place and the man, and of these the third is the chief variable. It is a pregnant thought that the mental make-up of every investigator has some special quality of its own, either inherent or due to past experience, an individual way of looking at things which will, after some manner, lead to discovery. However often a particular geological circumstance may have been considered by others, a new mind often makes progress which has been denied to its predecessors. But the capacity in the human mind of creating 'newness' is not in itself enough. I would say that two other fundamental qualities are required for all research. One of these is the power of self-criticism, and here it seems fitting to quote a remark made by Faraday: 'The world little knows how many of the thoughts and theories which have passed through the mind of a scientific investigator have been crushed in silence and secrecy by his own severe criticism and adverse examination.' Perhaps in no science is the need for self-criticism so insistent as in geology. In the field and laboratory the geologist is constantly being faced with problems of the most exacting kind and situations in which he must deal drastically with preconceptions of his own or of others, if these are not to prove his masters. With this quality the second is closely linked, namely, the desire for truth, at all costs. So, finally, I should like to recall a verse written by Arthur Hugh Clough, which I found copied out in a notebook of a relation of his, the late Dr. C. T. Clough, whose single-minded devotion in the pursuit of truth for its own sake was the most outstanding quality of one of the finest geological mappers the world has had.

'It fortifies my soul to know
That, though I perish, Truth is so;
That howsoe'er I stray and range,
Whate'er I do, Thou dost not change.
I steadier step when I recall
That, if I slip, Thou dost not fall.' [1]

[1] Delivered Sept. 8, 1952.

SOME PROBLEMS OF PARTHENOGENESIS

Address by

PROF. A. D. PEACOCK

PRESIDENT OF SECTION D

IT IS an honour to be your President ; for this mark of your goodwill, of which I am acutely sensible, I return my best thanks. In part-service lies this Address, the subject of which I chose—not selfishly, I hope—because it has interested me for thirty years. Nevertheless, I feel sure it also has much of interest for the layman ; certainly it has claims for much consideration by biologists, in their teaching and research alike. I therefore take this occasion, as one object of my Address, to indicate promising lines of work, to apprentice and journeymen biologists particularly.

The subject is wide, this gathering (without disrespect), heterogeneous ; these constrain me to present a mere sketch non-technically, thus risking over-simplification. Also, time is limited ; so once for all, I raise the lecturer's *cri du cœur* ' Time does not permit.'

Parthenogenesis, virgin birth, means the production of offspring from a reproductive cell, usually the female gamete, without the concurrence of a gamete of the opposite sex. The definition is carefully framed because male parthenogenesis occurs in certain algae whose male and female gametes differ little from one another, in contradistinction to the extremes seen in the minute spermatozoon and the large egg of animals.

Female parthenogenesis is of two distinct types, one resulting in male-production (arrhenotoky) and the other, in general, female-production (thelytoky). Both sexes are produced in certain cases (deuterotoky, amphotoky) but though this appears exceptional the explanation (see later) does not invalidate the general rule. To simplify terminology in this context the terms arrhenotokous and thelytokous are used as implying respectively male- and female-producing parthenogenesis, and the terms parthenogonic and parthenogone for parthenogenetically produced and an organism so produced.

Arrhenotoky occurs in wheel animalcules, saw-flies, bees and their allies, thrips, mites, scale insects and a Micromalthus beetle. The males are impaternate and produced unisexually, instead of bisexually as is general in animals. (Hence the caption in a New York paper, referring to a note on saw-flies I read in 1924 to Section D in Canada : ' Pity poor saw-flies; have mothers but no fathers.') Thelytoky reaches its complete expression when it is the exclusive, obligatory method of reproduction as exploited notably among rotifers, water-fleas and insects.

The true nature of parthenogenesis can only be appraised by study of parthenogenetic and bisexually produced eggs. The nuclei of ripe sex cells (gametes), by comparison with those of the body (somatic) cells, contain only half the number of chromosomes, bodies that are the main physical basis of heredity. This is a provision whereby at fertilisation the gametes with the reduced or ' haploid ' chromosome number co-operate to produce an organism whose cells possess the double or ' diploid ' chromosome number. An unfertilised haploid egg of a bisexual species normally cannot develop because viability depends upon a certain physiological relation between the nucleus and the surrounding cytoplasm. This relation, designated K/P, exists when the chromosome number is diploid. However, certain parthenogenetic haploid eggs are viable, in arrhenotoky straightway producing haploid males, but in thelytoky, by a special process of doubling the chromosome number, diploid females. This ' regulation ' of chromosome number

needs description. The nucleus of the mother cell of the egg contains two haploid sets of chromosomes, one of maternal and the other of paternal origin, each chromosome thus having a counterpart (homologue). In the ripening of the egg (maturation) the cytoplasm of the cell is divided into two unequal cells, one the second oocyte, the other the small polar body; further, the partner chromosomes separate so that each of the two cells contains a haploid set of homologous chromosomes. At a certain stage the chromosomes are circularly arranged as if on a plate. Next, each of these cells divides and each of the chromosomes splits into two identical daughters which part company so that one haploid set is contained in the ripe egg cell and in each of the three polar bodies. This series of events is called meiosis. Normally the unfertilised haploid egg dies. In thelytoky, however, the egg, or in some species the haploid cells of the organism developing from it, have somehow become endowed with the power of retaining or restoring the diploid number of chromosomes. In the first case, according to species, the first reduction division has vanished or one of the plates of polar-body chromosomes is retained; in the second case the chromosomes of the haploid set each reproduce themselves, split into two daughters, and all remain in the cell. The upshot, in any event, is that diploidy is obtained.

Some organisms in the course of evolution have, however, acquired multiple numbers of chromosome sets, and are polyploid. For example, in tetraploid cases the gametes contain two sets of chromosomes and are not therefore haploid; similarly for 'even-number' polyploidy the number of gametic sets is half the somatic number: maturation, however, remains the same. The special cases of odd-number polyploids need not be considered here.

In regulation nature is a conjurer performing vanishing and balancing tricks with plates of chromosomes.

This Address centres round three major themes: cyclic parthenogenesis, heterogony, the alternation of the parthenogenetic and bisexual modes of reproduction; sex determination; and artificial parthenogenesis. But first, however, the increasing toll of new cases of parthenogenesis reported merits some attention.

The only main groups in which parthenogenetic phenomena are now unknown are protozoa, sponges, polyzoans, lamp-shells, sea squirts and reptiles, for recently these phenomena have been discovered in earthworms, embiids, book-lice, beetles, moth-winged flies and spiders, all previously unknown or dubiously known as being parthenogenetic.

For earthworms, Gavrilov (1939) reported that a species of Eiseniella was parthenogenetic, a conclusion which Muldal (1952) in England has enormously extended. Obligatory parthenogenesis is reported from five genera which show triploidy, tetraploidy and decaploidy, and all the polyploids are larger than their nearest diploid relatives. Facultative parthenogenesis was also found in a tetraploid species. These cytological observations mark one of the greatest advances in the study of earthworms since the publication, seventy years ago, of Darwin's famous treatise.

Thelytoky is reported for the first time among moth-winged flies, in *Psychoda severini* var. *parthenogenetica*, by Tonnoir (1920).

Geographical parthenogenesis, so successfully studied by Vandel (1931, 1940), refers to the co-existence within a species or genus of respectively two races or species, one of which reproduces bisexually and the other parthenogenetically, the two groups occupying different habitats and differing cytologically, the parthenogenetic form often being polyploid by comparison with the bisexual. Interest in this increases. The bag moth *Solenobia triquetrella*, intensively studied by Seiler (1949), comprises three distinct races: (1) diploid and bisexual, (2) diploid and parthenogenetic, (3) tetraploid and parthenogenetic, and there seems no doubt that the first gave rise to the second and the second to the third. In Switzerland the first and second races are restricted to ice-free pockets bordering the glaciers but in regions which the Quaternary glaciers did not cover. The tetraploid race, however, is found all along the Alps and in Central Europe, and is evidently of recent origin and of post-Quaternary distribution. Its success in adapting itself

to hard conditions is doubtless due to its polyploidy, as in other cases illustrating geographical parthenogenesis. A truly remarkable coincidence is seen in Suomalainen's (1940) findings on weevils, several species of which exist in two races, one diploid and bisexual, the other polyploid and parthenogenetic. In Alpine species the bisexual race is limited to the comparatively warm valleys which surround the mountains and were never under Quaternary ice whereas the parthenogenetic race has occupied the central Alpine region. Seiler and Suomalainen, due to the War, were unaware of each other's work, a fact intensifying the interest of their discoveries.

The East Indian wood-louse *Nagara modesta*, as Hill (1948) shows, exists as two races on different islands, one form bisexual, the other thelytokous and tetraploid. The grasshopper *Sago pedo* studied by Matthey (1948) shows similar phenomena.

Paedogenesis, parthenogenetic reproduction by a pre-adult, has long been known in gall midges. Their cytology is remarkable because chromosome elimination occurs during development (see Reitberger, 1940 and White, 1945). The germ line of both sexes of *Miastor metraloas* contains 48 chromosomes but 12 and 6 only in the female and male soma respectively (White). During development 36 chromosomes are eliminated from the somatic nuclei and females develop ; and 42 when the males of a sexual generation develop.

Facultative or accidental parthenogenesis (tychoparthenogenesis), by which diploid parthenogones are produced in a diploid bisexual species, plays no part in the reproduction of that species. Another specialisation, recorded by Swann and Mickey (1947), is, however, seen in the facultative thelytoky of the North American grasshopper *Romalea microptera*, the tissues being predominantly haploid. Here may be noted Fankhauser's (1938) rearing of a haploid female newt by artificial parthenogenesis. Haploidy in bisexually-reproducing species does not then always imply nonviability.

The cytology of the honey bee, for forty years rather puzzling because of Nachtsheim's descriptions of clumped chromosomes, has been further resolved by my colleagues Sanderson and Hall (1948). They find no such chromosomes, and it would appear that the chromosome clumping described was due to the techniques employed. In general, bee gametogenesis is seen to follow that described for other hymenoptera such as saw-flies. A distinctive chromosome in germ cells of the male, two such in the female, were also noted but these workers (1951) desire much more information before agreeing that these bodies are sex chromosomes, as is the view of Manning and Kerr.

Rudimentary parthenogenesis in the gall wasp *Biorrhiza pallida*, from unpublished work by my colleague Dr. Ann Sanderson, is characterised in virgin eggs as follows : in the first few hours after laying, two groups of chromosomes are seen, as in the fertilised egg ; later, up to two or three days, up to ten large and polyploid nuclei develop, but thereafter degeneration supervenes.

Pseudogamy, in which the spermatozoon activates the egg but whose nucleus does not fuse with the female's, has been known in few cases. Now Lepori (1950) adds the case of the hermaphrodite flatworm *Polycelis nigra*, of which in Italy there appear to be di-, tetra- or octoploid forms. Hybridisation between the moths *Selenia bilunaria* (females) and *S. tetralunaria* males has resulted in Harrison (1933) obtaining a few pure *S. bilunaria* females, possibly a case of ' expedited ' parthenogenesis, as *S. bilunaria* eggs sometimes show rudimentary parthenogenesis.

Parthenogenetic merogony, the development of egg-fragments, has been induced in an Arbacia sea-urchin by the American worker Dr. Ethel Harvey (1936). Centrifugalized before the disappearance of the nuclear membrane, the eggs divided into spherical halves and quarters, and those without nuclei, activated by hypertonic sea-water, developed into blastulae (one with 500 cells) and, in one case, into an embryo sans gut, sans anus, sans skeleton. These findings raise fundamental issues. The cytoplasm is evidently of great importance in determining at least the early stages of development ; but whether its power is innate, or derived from material received from the nucleus previous to or

at the time of the breakdown of the nuclear membrane, is still an open question.

We come now to our main themes. Cyclic parthenogenesis, heterogony, connotes the alternation of bisexual and parthenogenetic modes of reproduction. The phenomenon is seen in wheel animalcules (Rotifera,) water-fleas (Crustacea Cladocera), plant-lice (Aphididea), gall gnats (Diptera Cecidomyidae) and gall wasps (Hymenoptera Cynipidae). In certain rotifers and water-fleas the cycle occurs once or more annually with such striking regularity that earlier observers, e.g., Weismann, considered the cycles obligatory and rhythmic, that is, determined solely by hereditary factors. This view is now discarded as experimentation shows environmental factors to be of primary importance. The problem today concerns the rôle of such factors. Until recently, experimental results showed such variety and wide discordance that no definite conclusions could be formulated, the explanation being the impossibility of comparing results obtained from a diversity of species subjected to a diversity of techniques themselves imperfect. The problem became soluble only when accurate knowledge of the heredity of the living material and of the operating environmental factors was available ; and this, in turn, had to await the development of refined techniques involving strains of organisms genetically pure and culture media suitable for the experimental animals and the micro-organisms forming their food.

Work on water-fleas by Banta (1939), initiated about 1914 in the U.S.A., deserves first mention. His culture medium and culture technique do not satisfy present standards but convincing results were nevertheless obtained. Constant parthenogenesis was maintained in *Daphnia pulex* for many years and through 767 generations, a family history far outrivalling that of the ' penniless lass wi' a lang pedigree.' No obligatory, rhythmic hereditary tendency therefore operated to override the effect of environmental factors and so induce bisexual reproduction. The induction of males and of bisexual reproduction could however be invoked in any generation of a long strain by the environmental factors coming into operation when parthenogenetic females were overcrowded. The operative factor, Banta considered, was the excess of excretory products accumulated in the culture bottles. Banta's explanation is not confirmed by recent workers but his conclusion that heterogony depends upon environmental factors is correct in the main.

Recent admirable researches, exploiting the most refined and satisfactory techniques, are those of German workers : Hans Buchner (1936) on rotifers and Madeleine von Dehn (1944) on water-fleas. Their results, combined with those of certain earlier workers, permit the following conclusions on rotifers and water-fleas. (When a group is not mentioned it is meant that information on a particular point is lacking for that group.)

1. *Continuous parthenogenesis* can be maintained. It is possibly the usual method of reproduction in certain heterogonous rotifers (Buchner).

2. *Co-existence of bisexual reproduction and heterogony* is regarded as obtaining in rotifers (Buchner) as a low percentage of males in control, stem and experimental cultures always occurred. Hence no complete switch-over from parthenogenesis to bisexual reproduction is likely.

3. *Environmental factors determine the cycle* in both rotifers and water-fleas, there being no hereditary mechanism operating in a fixed and rhythmic manner.

4. *Quality of food is the main factor known to induce bisexual reproduction* in rotifers and water-fleas. Light deficiency associated with food deficiency, and darkness also, are effective to a certain extent in water-fleas. No effects were obtained by temperature or excretory products in both rotifers and water-fleas. Further, quantity of food, light and air humidity were inoperative in rotifers, as was overcrowding in water-fleas.

Regarding quality of food Buchner comments that it must obviously vary because, according to their age and sexual condition, the food organisms must vary physiologically. Any type of waxing and waning of the operative agent would produce periods of strong or weak sexuality in the rotifers. The fatstuff content of food has recently been shown as important by von Dehn (1950).

5. *An inherent maximum potentiality for bisexuality* (Reaktionsnorm) is a physiological character fixed at a certain level for each race of a rotifier species and varying from race to race. Rhythmic variation of the strength of the norm may possibly occur though Buchner found only one rare case.

6. *The level of bisexuality attained* (Bisexualitätsgrad) in a line or race of rotifiers depends upon the strength of the environmental factors and of the factor determining the maximum potentiality for bisexuality, the former operating until a level is reached at which the latter imposes its effect.

7. *Obsolescence of the terms 'cyclic bisexuality' and 'alternation of generations' is the case with reference to rotifiers.* The former term must be abandoned because its original connotation, the fixed and rhythmic appearance of bisexual forms (mictics) no longer holds. The latter term has also lost its original meaning because, as a rule, generations having different sexual tendencies do not give rise to one another. What remains of its meaning is seen in the fact that the fertilised resting egg gives rise to a parthenogenetic generation.

Among gall gnats (Diptera Cecidomyidae) heterogony may co-exist with paedogenesis, and Ulrich (1934) finds in *Oligarces paradoxus* features similar to those recorded for rotifers and water-fleas. Continuous parthenogenesis by paedogenetic larvae for 263 generations has been maintained under constancy of environmental conditions, the bisexual adults appearing only when larvae were subjected to changes in the milieu involving light, temperature and crowding. No rhythmically-operating hereditary factors have so far been discovered.

In plant-lice (Aphididae) the rôles of hereditary and environmental factors require further investigation, but it would appear from Davidson's work (1929) on the bean aphis, *Aphis fabae*, that the latter are certainly important. In this species the sexuals and their offspring, the stem mothers (fundatrices), are found on a primary host in autumn and the diverse parthenogenetic forms of successive generations on an intermediate host in the spring and summer. Davidson used the spindle tree (*Euonymus europaeus*) as primary host and the broad bean (*Vicia faba*) as intermediate. By rearing one type, the alienicolae, of the parthenogenetic forms on broad beans grown all the year round in a greenhouse under suitable conditions of temperature and light, continuous parthenogenesis was maintained, the longest period being nearly three years and the number of generations 50. On reduction of the number of hours of daylight to which colonies were exposed from the fundatrix stage onwards, sexual females developed earlier than normally, in June, though males did not appear until the normal time, October. Lowering of temperature also induced sexuals out of season. The effect of nutrition, due possibly to the host-plant being affected by these changes in light and temperature, was not determined.

Whether an inherent fixed maximum potentiality for bisexuality exists, as in rotifers, is another problem in parthenogenesis. The general situation in plant-lice appears to be that heterogony is characteristic of species of temperate regions; that the more primitive the species the greater the influence of external factors, the more labile the cycle; that the more complicated the life-cycle the less this influence, until in forms with almost rigid life-cycles, e.g. *Phylloxera*, the vine pest, and the gall-forming *Chermes* of conifers, the appearance of sexuality is predominantly due to hereditary factors.

Regarding gall wasps no such investigations as described above have been made, but from certain similarities of habit it would not be surprising should conditions parallel those found in plant-lice.

We come next to our second main theme. Sex determination in arrhenotoky has long been a problem as it is unexplained by the theory of genic balance which holds for most bisexual organisms. This theory formulates that the sex of an organism depends upon conditions set up between two special (sex) chromosomes, 'X' and 'Y,' and the other chromosomes (autosomes), 'A', of the cell. The diploid constitution of the female is 2A + XX; that of the male 2A + XY, the Y chromosome being a homologue of an X and special to the male. Eggs are therefore all of the same constitution, A + X, but spermatozoa are of two kinds, A + X and

A + Y. Sex determination therefore depends upon whether the egg receives an X or a Y spermatozoon. In thelytoky, the general rule being that the eggs are all of one kind, only females arise by parthenogenesis. Where both sexes in a diploid bisexual species are produced by parthenogenesis (deuterotoky), as in moths, it is the female that produces gametes of two kinds and, therefore, both male and female parthenogones. In arrhenotoky mothers of constitution 2A + XX produce haploid parthenogones of constitution A + X; the genic balance is undisturbed so that one would expect female offspring instead of, as they are, male. The riddle was solved by the Whitings in the U.S.A. for the little parasitic wasp-like ichneumon *Habrobracon juglandis*, an important clue being the discovery, about thirty years ago, of rare diploid males (Whiting, 1945), arising when closely related stocks were interbred. Their spermatogenesis is similar to that of haploid males, their spermatozoa are diploid and, genetically, they do not behave as normal diploid organisms but, we may say, as di-haploids. They also show that maleness in arrhenotoky is not necessarily a consequence of haploidy.

Whiting's multiple-allele theory of sex determination in hymenoptera is based on the genetics of normal males showing certain mutant traits, diploid males and normal females. The results of ingenious breeding experiments justify the following inferences. A sex segment, *x*, lies in a certain chromosome and exists in a number, nine so far known, of slightly modified forms (alleles), which may be designated *xa*, *xb*, *xc* . . . *xi*. This sex segment comprises many dominant female-producing genes and recessive male-producing genes responsible for the numerous sex differences. The sex of an individual depends upon which alleles are present, that is, on their complementary action. Dominant and recessive genes also lie on either side of the sex segment. A female, being diploid, possesses one each of two such different alleles, *xa*/*xb* or *xb*/*xc*, etc. and is heterozygous; its parthenogonic male offspring, being haploid, can be expressed *xa* or *xb* or *xc*, etc.; and the rare diploid male, the offspring of a pairing, is *xa*/*xa* or *xb*/*xb* or *xc*/*xc*, etc., making it homozygous. Thus in a pairing between an *xa*/*xb* female and an *xa* male the female parthenogenetically would produce individuals expressible as *xa* and *xb*—males, and bisexually, individuals *xa*/*xa*—diploid males (rarely viable) and individuals *xa*/*xb*—females.

In the different hereditary types (genotypes) females are heterozygous dominants for many of the sex genes; conversely, the rare diploid males are homozygous and the functional haploid males azygous, both showing more of the recessive traits determined by the male-producing genes.

We may regard the Whiting theory as one of the greatest steps in hymenopterous studies since the time of the great bee-master, Pastor Dzierzon, who showed in 1845 that in the honey bee unfertilised eggs give rise to males and fertilised to queens and workers. (See v. Siebold, 1857.) Further it is the most satisfactory explanation to date of sex determination in other arrhenotokous groups.

Another outstanding example of the impact of parthenogenesis research on general biological problems is seen in the results of the Seiler school in Switzerland on intersexuality and sex in Solenobia moths. This work must be seen against the background of Goldschmidt's (1923) brilliant researches on the Gipsy Moth, *Lymantria dispar*. Throughout Europe and Asia there exist different races of this moth and according to the type of racial crossing Goldschmidt obtained intersexual offspring, ranging from females slightly masculinised to males slightly feminised. Racial inbreeding produces normal offspring, which, arguing from the obvious fact that intersexes carry the sex-determining substances of both sexes, implies that normal organisms of both sexes also carry the sex-determining substances of both sexes in harmonious relationship. If F represents the female-determining and M the male-determining substances the female, with F substances predominating, can be symbolised as $F < M$, and, conversely, the male as $F > M$. An intersex results when the normal quantitative relationship between F and M is disturbed, such as occurs when the male and female gametes pertain to different races. From the study of the abnormal, Goldschmidt thus explains the

normal. Regarding intersexuality Goldschmidt formulated his 'Zeitgesetz' or time law. This will be considered along with the Seiler findings.

From Seiler's labours (1949) on *Solenobia triquetrella* representative evidence on intersexes comes from crossings between females of the tetraploid parthenogenetic race and diploid males of the bisexual race. Now one of the excellences of this work is that values have been assigned to these intersexes by measuring the maleness and femaleness observable in many of their organs. Crossings involving these females from different localities gave all grades of intersexuals; the grade differed from crossing to crossing, evidence that the tetraploid parthenogenetic race comprises geographical sub-races. A most remarkable feature, however, was that from certain pairings all grades of intersexes were found in individual broods and not, as might be expected, one-grade intersexes. From evaluations of *individual* intersexes three critical points arise: (1) all organs in general exhibit the same grade of intersexuality; (2) the same grade of intersexuality exists in pupa and imago; (3) no wild disorder of grades of intersexuality exists within an intersex at any stage and therefore there is strong probability that the grade of intersexuality, and the sex, of an individual moth is determined possibly in its embryonic, at the latest in its larval, stage. Seiler's important interpretations and their points of difference with Goldschmidt's may now be cited:

1. *Time and mode of action of the sex factors.* Both F and M operate simultaneously right throughout the development of the intersex and, contrary to Goldschmidt, not successively and by one gaining ascendancy over the other.

2. *Grade of intersexuality.* Each intersex begins and continues at the same grade; there is no switch-over from one sex to the other as Goldschmidt supposes.

3. *The nature of an intersex.* An intersex is a mosaic of female and male organs all of which, in general, possess the same grade of intersexuality right from the beginning of their development; it is not, contrary to Goldschmidt, a mosaic comprised of the early-developing organs of one sex and later-developing organs of the other sex.

4. *How different broods come to differ in their grade of intersexuality.* This occurs after crossings of males of the same origin and hereditary (genetic) nature with tetraploid parthenogenetic females of differing origin. Some broods obtained comprise female-like intersexes, others comprise male-like intersexes, so that obviously the parthenogenetic female parents of differing origins must have differed genetically. This recalls Goldschmidt's findings of genetically different races of Gipsy Moth.

5. *How all grades of intersex come to be found in the same brood.* This is a crucial discovery. The genetical constitution of the hybridising individuals was known to be fixed, so that if genetical factors alone operated all the offspring in each individual brood would be hereditarily identical. As, however, all grades of intersexes arose, the only explanation left is that other factors, called phenotypical, came into play and modified the effects of the hereditary factors. In other words, we see an example of how the outward appearance of an organism (the phenotypical condition) does not necessarily reflect the hereditary make-up (the genotypical condition). The nature and modes of action of these phenotypical factors are unresolved problems.

Seiler's position may be summarised thus. Goldschmidt's theory of sex determination is held to be correct, as is the view that the different grades of intersexuality in the Gipsy Moth are conditioned *essentially* by hereditary factors. But as regards intersexuality, Goldschmidt's 'time-law' interpretation is held invalid, as it fails to explain such Solenobia findings as the occurrence in individual broods of intersexes of varying grade and the uniform grade of intersexuality throughout the organs of individual intersexes. In an intersex the F and M factors are in a certain fixed relationship throughout its life-history.

We turn now to researches on artificial parthenogenesis. A natural development of these, one might expect, would have been the establishment of parthenogenetic strains derived from artificial parthenogones. Such has not occurred

the best explanation probably being that, like the expert photographer on obtaining a first successful print, researchers simply lost interest in the original matter. Once a new thing is disclosed, a veritable diaspora occurs among its investigators ; each, according to need or fancy, goes chasing his own hare—and rightly so. Hence, out of the discovery of artificial parthenogenesis, arose new ideas on reproduction, cytology, cell physiology, biophysics and biochemistry.

The solitary example of a parthenogenetic strain stemming from artificial parthenogones is recorded by the Russian worker Astaurov (1936) for the silk moth, three successive parthenogonic generations being obtained. The activating technique consisted of subjecting unfertilised ovarian or newly-laid eggs to increased temperature.

The next most successful results are limited to young frog parthenogones advanced enough to be sexed. In his early pioneer work Loeb obtained by needle traumatisation 46 metamorphosed tadpoles and another 68 tadpoles. Of a batch of 9 metamorphs 7 were male and 2 female ; in a second batch of 34 well-developed tadpoles (including 1 metamorph) 12 were male, 2 were changing towards maleness, 18 were female and 2 were doubtful ; in a third batch of 21, 3 could be recognised as female. Later needle traumatisation work, 1933, by Parmenter (1940) yielded 26 tadpoles, of which 5 metamorphosed, the ages of the parthenogones varying from three to seventeen months. The explanation of the production of both sexes is too complicated for discussion here.

By cytological standards a good proportion of these parthenogones were normal, for they were diploid : 15 metamorphs and 35 tadpoles (Loeb) ; 5 metamorphs and 4 tadpoles (Parmenter). The range of cytological conditions disclosed by Parmenter in 26 frog tadpoles was remarkable ; haploids, 10 ; diploids, 9, of which 5 metamorphosed ; haploid-diploid forms, 4 ; diploid-triploid forms, 1 ; diploid-triploid-tetraploid forms, 1 ; and triploids, 1.

A few young adults, unsexed, have also been obtained in a starfish, sea-urchins, and a fish. Regarding the starfish *Asterias glacialis*, studied by Delage (1902), there is evidence from Paul Buchner (1911) that parthenogones were diploid and normal. Parthenogones of the sea-urchins *Strongylocentrotus lividus*, obtained by Delage (1913), and *Echinus esculentus*, obtained by Shearer and Lloyd (see Vandel, 1931), from our present small amount of knowledge are haploid, and it is interesting that some of the earliest evidence came from Dr. Hindle (1911), till recently a General Secretary of the British Association, then a junior colleague of Jacques Loeb. By a ' bizarre ' experiment, in which carp eggs were stirred in human saliva, Kasansky (1935) has obtained adults 8½ months old, normal to all appearance.

The greater success attending the silk moth experiments is probably due to the insect's terrestrial habit : the larval and later stages are much more easily reared as their milieu, by comparison with the aquatic, is much less complicated and susceptible of change, and the problem of their feeding is simpler.

Loeb (1913) in the U.S.A., in 1899 was the first to obtain advanced larvae from artificially activated sea-urchin eggs, using magnesium chloride in sea-water. Likewise, Bataillon (1910) in France, in 1910, was the first to obtain living tadpoles by pricking frog eggs with a blood- or lymph-soiled needle, reviewed under the title ' The new Aaron's rod ' by a former President of this Section. We all know (or do we ?) that the rod of Aaron budded, but how many remember that the rod in Aaron's hand was stretched out over the waters of Egypt ' and the frogs came up and covered the land of Egypt.' The reviewer ? Who but D'Arcy Thompson.

Since Loeb's success there have been exploited at least 371 methods—I have counted them—45 physical, 93 chemical, 64 biological and 169 combined, not to speak of variations innumerable. Every agent common in the physics, chemical, or biological laboratories seems to have been tried, as well as onion juice and mustard from the kitchen. The animal groups and species in which parthenogenetic phenomena have been induced number about 9 and 70 respectively. With the exception of insects, ticks and mammals,

all are aquatic and comprise ribbon worms, bristle worms, molluscs, starfish, sea-urchins and fish, as well as freshwater fish and amphibians.

The cardinal question concerning natural fertilisation and parthenogenesis is activation. In other words, what is the process that sets going that most marvellous thing in nature, the development of a new living thing. It was natural that investigations on phenomena revealed in the egg by artificial treatments and on those revealed by sperm fertilisation should become complementary. Living material has been obtained principally from marine worms, frogs, and *par excellence* from Arbacia sea-urchins, which are singularly favourable, being plentiful and prolific of gametes lending themselves easily to experiment. This came home to me recently when I was privileged to occupy the British Association table at the Stazione Zoologica, Naples ; of some forty researchers there from several nations a third were doing things to or with Arbacias. Arbacia is assuredly the marine experimenter's Drosophila.

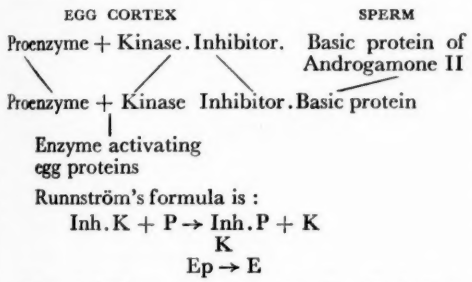
Despite the work of half a century, activation still remains an enigma. However, the labours of R. S. Lillie, Runnström, Tyler and others (see Tyler, 1941 ; also Munroy, 1950), have shorn the problem of side issues so that investigations can now be concentrated more on the surface layer (cortex) of the egg, and not on the outer layers which, essentially, are ancillary specialisations concerned with protecting the egg and with fertilisation phenomena occurring outside the egg proper. The outermost layer, the egg jelly, proves to be concerned with the activation, attraction, and penetration of the sperm, and with the exclusion of superfluous sperms. The layer within, the vitelline membrane, after insemination of the egg, becomes the fertilisation membrane when incorporated with granules secreted and freed from the cortex. Inner again is the cytoplasmic or hyaline membrane, so little studied that it invites research. The cortex is the first living part to encounter environmental stimuli.

Activation being polymorphic, two possibilities arise regarding the ' touch-off ' of the series of processes constituting development : the egg may possess either a number of critical ' touch-off ' substances, only one of which reacts to a particular agent or agents, or only one such substance sensitive to all agents. The problems here implied await solution. On activation there occur in the egg many physical and chemical changes, but most of these need not be considered here as they are mainly associated with the metabolism of development. Whether activation is solely due to cortical processes is by no means certain though all investigation indicates their major importance. But some idea of cortical events, at least for sea-urchin eggs, can be gained by considering observations on natural and artificial activation. The idea sketched here is based on recent information gained from the papers and conversation of researchers at Naples, but current investigations proceed so vigorously that it must be regarded as very provisional.

The cortex is mostly composed of fats (lipids) and at the touch of the sperm or on artificial treatment breakdown changes (cytolysis) and liquefaction occur. Polariscope observations (see Munroy and Montalenti, 1947) prove the molecular disarrangement of the cortex, the agent responsible being a known sperm acid similar to one causing the breakdown of blood (haemolytic). This acid is possibly an important component of the sperm substance designated Androgamone III. That the sperm produces this effect chemically is borne out by Kling's observations (1936) that, for frog eggs, a needle of the same diameter as a sperm is not so activating in function as a needle of greater diameter. The cortex recovers, as the polariscope proves, the explanation offered being that the sperm carries a corrective to its cytolytic agent. The processes whereby superfluous sperms are excluded and the fertilisation membrane produced also cause cortical changes, but these are not regarded as directly affecting activation. The critical changes involve permeability, whereby substances are enabled to pass in either direction through the cortex, causing alternate egg shrinkage and swelling, resulting in rearrangements of, and reactions between, materials in the egg substance. One

important cause of change in the composition and arrangement of the cortex, rendering it permeable, is believed to be the loss of calcium and, possibly, other electrolytes.

What emerges is Runnström's (1949) simple hypothesis of activation by ferment action. The unactivated egg cortex is regarded as containing an inactive proenzyme and a kinase whose activity is held in check by an inhibitor, a glycoprotein known to prevent the action of protein-splitting ferments. The sperm, among several other substances of different functions, carries to the egg Androgamone II, containing Antifertilizin (Tyler) and a basic protein (Runnström). On insemination the basic protein binds the inhibitor substance of the egg, thus freeing the kinase. The free kinase activates the proenzyme into an enzyme, which, acting on the egg proteins, initiates development. This can be expressed as follows :



Runnström's formula is :

$$\text{Inh.K} + \text{P} \rightarrow \text{Inh.P} + \text{K}$$

$$\text{Ep} \xrightarrow{\text{K}} \text{E}$$

Inh = Inhibitor ; K = Kinase ; P = Basic Protein ; Ep = Proenzyme.

The basic protein and the enzyme mentioned are known to occur though their chemical formulae and those of the other substances mentioned are unknown. But this hypothesis, there is reason to believe, may only relate to a phase of activation following another involving a sulphydride (SH) group.

Figuratively, the egg, activation and development resemble a factory and its activities. In the factory some agent throws a switch over, energy is released, wheels turn, belts move, machines take in stuffs to move them about or break them up or re-combine them variously. Here the simile ends. For, regarding the egg, we need to imagine it as a factory whose very framework, protein, is alive ; whose mechanisms are so minuscule, intricate, delicate and precise as to transcend anything graven by art and man's device ; whose materials are magically labile, mobile, dissolving, interacting, and changing ; whose climax is self-cleavage into two factories, themselves fissile.

Despite the many difficulties inherent in the living material, experimenters have not been deterred from attempting to obtain parthenogones from mammalian eggs, and astonishing results have been obtained at the hands of Pincus (1936, 1951a) and co-workers (1951b) in the U.S.A., using rabbits, and of Thibault (1948–49) in France, using rabbits, rats and sheep. Specialised techniques of various workers have been exploited. A plentiful supply of eggs from mammals normally producing few at a time was obtained by injecting a pituitary hormone which increased ovulation, e.g., 100 and more eggs at one time being produced in the rabbit. To keep eggs alive outside the animal they could be stored for several days at refrigerator temperatures. A medium of blood serum in watch-glasses served for cultures. It being at present impossible to rear such cultures for more than a few days, the eggs had to be transplanted into a host animal. The latter should be in the same physiological state as exists in the initial stages of pregnancy and this state, called pseudo-pregnancy, is induced by administering a pituitary hormone or by mating with a male operatively treated so that spermatozoa cannot be emitted. Yet another technique is cold treatment of eggs in the Fallopian tube, the tube being exposed by dissection and enveloped in a metal cylinder through which cold tap-water circulates.

The most extraordinary achievement of the Pincus school (1951b) has been the development of seven rabbit parthenogones, all female, one of which produced a litter of nine after mating. Five of the parthenogones were born alive and one dead, the seventh went to full term but was never born. The living comprised three adults and two young. Hypertonic treatment (by common salt or other saline solution) of oocytes and their subsequent transference to pseudo-pregnant foster mothers accounted for five animals—

three living adults, one stillborn young, one young in uterus ; watch-glass culture and *in vivo* cooling treatments produced one young each. Pincus draws attention to the very low proportion of living parthenogones to oocytes artificially treated, about 1 : 2500, and how unfavourably it compares with the successes, up to 100 per cent., obtained after similar treatments of sea-urchin eggs.

Of the less successful results with rabbit eggs in the ovary and in the Fallopian tube mention need be made of but few. Using tubal oocytes subjected to cold *in situ* Thibault obtained a high activation rate and a few blastocysts, one of which effected fixation in the uterus. Similar results eventuated when cold was applied to oocytes by the metal-cylinder technique and a still higher activation rate by the application of ice. Thibault also used the tubal oocytes of sheep and rat. In the sheep activation and cleavage were achieved *in situ* using cold (ice for three minutes), ' un refroidissement brutal et bref.' Activation, 100 per cent., was also observed after similar cold treatment and after barbyturic or ether anaesthesia caused fall of body temperature, observations causing Thibault to sound the warning that in interpreting results the possible effects of the anaesthetics should be borne in mind.

The cytology of these cases has been investigated by Thibault, following on preliminary work by Pincus. Regulation in the tubal oocytes of both rabbit and sheep occurs at the second maturation division, when the chromosomes remain in the diploid number. No such regulation occurs in the rat ; maturation runs full course. There is much scope, however, for further cytological research.

Regarding the evolution of parthenogenesis let us first consider arrhenotoky and, arising from it, the issues of the production of haploid spermatozoa, of male haploidy, and of haploid-male viability. At the outset we encounter the difficulty as to whether animals regarded as haploids are really so. For recent research on polyploidy emphasises that the idea of diploidy being characteristic of somatic cells and haploidy of germ cells requires revision. What is considered as the haploid number may be some multiple of an ancestral basic number (the ' monoploid ' of Whiting). In other words modern haploid and diploid animals may be polyploid by comparison with their remote ancestors and the terms ' haploid ' and ' diploid ' be relative. If this were so it would remove the difficulty of how a suitable relation between chromosomes and cytoplasm (K/P) exists in so-called haploids. However, the following facts concerning the singular characteristics of males resulting from arrhenotoky lead to an alternative conclusion.

The maturation of the male gamete in a normal diploid male parallels that of the egg except that four functional haploid gametes are formed. Now in arrhenotokous species spermatogenesis is peculiar in that the reduction division is abortive, so that the number of chromosomes in each of the two gametes produced remains the same as in the sperm mother cell. It is obvious that reduction would result in sub-haploidy of germ cells and would defeat the purpose of fertilisation so that this regulation of chromosome number is an adaptation associated with the evolution of haploid males. How the reduced condition and the viability of males have originated and evolved are suggested from the work on scale insects by the Schradens (1931) in the U.S.A. Haploid males are regarded as modifications of diploid males of an originally normal bisexual species, as the result of progressive degeneration in one haploid set of chromosomes. At one end of a series a certain species of scale insect produces normal diploid sperms by normal spermatogenesis ; at the other end, in a different species, an entire haploid set has disappeared and the functional haploid set is found in the haploid sperms that are produced in only half the normal number ; in between, different species show different conditions of such degeneration. Further, the males of one species even show a degenerating haploid set in the body cells and thus are in effect haploid. Although these phenomena do not occur in species of a single line of descent they can be interpreted as illustrating the different stages of degeneracy through which the males of a given single species passed in their progress from diploidy to haploidy ; and, further, that viability has depended upon the gradual

manner of this degeneracy and not upon any parthenogenetic capacity acquired by the eggs from which the males originated.

This interpretation is challenged by Whiting (1945), principally on the ground that the cytological abnormalities seen in the scale insects are seen also in non-parthenogenetic organisms, and he offers alternatively his race-hybrid hypothesis. He suggests that the material which would contain the necessary genes for parthenogenesis and which would gradually provide the viable haploid male parthenogones could arise from hybridisation initially between males of a normal bisexual race and females of a parthenogenetic race, later between males and hybrid females, and later still between members of the segregating fraternities.

With these previous questions runs that of the nucleoplasmic relation in haploid cells. Explanations remain speculative. The appropriate K/P relationship could be established either (1) by diminution in the amount of cytoplasm with reduction in the size of the egg and the male ; or (2) by increase in chromosome size. The first may conceivably be the case in rotifers where the fatherless haploid eggs (and the males arising therefrom) are smaller than those produced bisexually. The second alternative is favoured by Whiting (1945) from, *inter alia*, observations on cell-size in diploid males and on the genetics of males of Habrobracon.

That parthenogenetic forms have originated from bisexual forms is evident. Parthenogenetic females of numerous groups still possess a spermatheca, an organ which in normal bisexual females of these groups stores the sperms received. Morphological resemblances of polyploid to diploid races are seen in brine shrimps (Artom, 1931), wood-lice (Vandel), moths (Seiler) and weevils (Suomalainen) ; further, in the wood-lice and moths mentioned, the diploid bisexual males can pair with the polyploid parthenogenetic forms. The abrupt appearance of an exclusively female strain in a bisexual culture of nematode worms has actually been observed by Paula Hertwig (1920). The mutants were not parthenogenetic but were peculiar in that egg development depended solely upon sperm activation without the co-operation of the sperm nucleus. This case demonstrates a stage in the evolution of parthenogenesis. The next stage, at which the egg acquires the further faculty of dispensing with sperm activation, has never been observed.

According to the type of egg, regulation can occur either after maturation (during development), or at different stages during maturation. Now, just as in the acquirement of any new character, the faculty of regulation has been acquired as the result of permanent change in the hereditary material, that is, of mutation in the gene(s) governing the relevant chromosome division. The new genes might act at any stage. If at first they acted after maturation, and if in the course of evolution the genes operated earlier and earlier, the stage at which regulation occurred would likewise be progressively earlier in the development and maturation of the egg. Such a course would make thelytokous parthenogenesis haploid in origin. On the other hand, in experimental parthenogenesis, where the same type of egg or different types of eggs can be made to regulate at different stages, results clearly indicate the possibility that for different eggs mutations affecting different stages of maturation has occurred.

Unfertilised eggs which can develop only so far exhibit rudimentary parthenogenesis ; this occurs in most great groups from worms to mammals. Further work will undoubtedly disclose many more cases. Non-viability is attributed to the facts that the developing egg remains haploid, has an unsuitable K/P condition and has not acquired the faculty of regulation.

A further stage in the progress towards constant parthenogenesis is seen in facultative or accidental parthenogenesis (tychoparthenogenesis). The numbers of groups and species involved are somewhat limited, but further work will very probably add to the list, which comprises insects—grasshoppers, scale insects, moths (including the silk moth), and echinoderms—starfish and sea-urchins. Here the ripe egg begins with the reduced chromosome number, but during development the organism acquires the full constitution (diploid or polyploid) in its cells by various methods of regulation similar to those occurring in certain cases of both obligatory and artificial parthenogenesis.

The association of polyploidy and parthenogenesis in several species of animals is not regarded as generally one of cause and effect; rather, by becoming polyploid parthenogenetic forms have become enabled to people inhospitable regions. The idea that the scarcity or disappearance of males is the cause of parthenogenesis is discredited. It is a case of parthenogenesis, once acquired, rendering males unnecessary, and in the inevitable course of nature, their final disappearance.

Our understanding of the evolution of parthenogenesis will become clearer only by further research. In this connection researches on the viability of haploid parthenogones as seen already in the cases of a haploid female grasshopper and a haploid female newt, will prove of much importance.

The factors in nature that have induced parthenogenesis are completely unknown, though experiment shows physical and chemical changes in the milieu to be effective. Hybridisation of plants in nature, e.g., wild roses and hawkweeds, is regarded as significant by botanists; but for animal parthenogenesis what little is known relates mainly to crossings between two non-parthenogenetic moths, *Tephrosia crepuscularia* and *Tephrosia bistortata*, as the result of which Peacock and Harrison (1926) recorded hybrid offspring capable of parthenogenetic reproduction. Except for these hints, the origin of parthenogenesis remains an enigma.

Parthenogenesis does not preclude the evolution of new forms. For example, certain water-fleas in nature exist as local races, each with special hereditary characteristics. Again, in parthenogenetic laboratory strains mutations and their inheritance have actually been witnessed by Banta (1939). The novelties showed 'excavated head,' 'short beak,' sex intergradedness, great reactivity to light, and the need for higher temperature. Local races of rotifers, differing morphologically and in their potentiality for producing sexuals, are also known. It is probable that among the specialised Chermes plant-lice of conifers certain distinctive forms of individuals of the parthenogenetic phase have arisen parthenogenetically from a prior type. In the arrhenotokous Habrobracon numerous mutants have been discovered by Whiting.

In parthenogenetic races we see species in the making. Such races can have originated in three main ways. First, in a bisexual species two types of parthenogenesis have arisen, resulting in the appearance of two different kinds of parthenogenetic races: e.g., an arrhenotokous race later supplanted by a thelytokous race, as in certain saw-flies; or else a thelytokous race and a tychoparthenogenetic race, as in certain scale insects. Secondly, a heterogonous species, migratory between two host plants, becomes a purely parthenogenetic form if one host plant is lacking; examples occur among green-fly and gall wasps (see Vandel, 1931). Thirdly, a bisexual species gives rise directly or indirectly to a parthenogenetic polyploid race, as seen in geographical parthenogenesis. That these new forms are actually distinct is supported by evidence that, as a rule, offspring from pairings with the parent type, when such pairings occur, are not produced or are abnormal. These distinct forms, then, especially those occupying new ecological regions, are free to evolve along their own lines into species.

Despite all the aforementioned curious and interesting things about parthenogenesis its biological significance should not be over-stressed. It must be remembered that the most characteristic method of animal reproduction is bisexual. Certainly parthenogenesis is very widespread but as the sole method of reproduction its extent is generally limited. For instance, it is of no account in the vast group of vertebrates while among invertebrates it is chiefly found among rotifers and insects. Its prevalence among insects is scarcely surprising, for insects comprise about half the known animal species. Within this class the many groups of hymenopterons, plant-lice and thrips (probably) have exploited thelytoky most fully. As a method of male-production it is characteristic of (1) the arrhenotokous Hymenoptera, including the heterogonous gall wasps; Hemiptera—white flies and scale insects; thrips, so far as known; heterogonous rotifers most probably; mites and a Micromalthus beetle; and (2) the heterogonous water-fleas and plant-lice.

The adaptive value of parthenogenesis lies in the fact that more rapid reproduction is possible than by the normal bisexual mode. Lack of males means more females in a generation and more females for the reproduction of the next generation. Further, as well expressed by Banta (1939) : 'An advantageous combination of characters or a favourable mutation once it occurred, barring accident, should continue and come to prevail in the habitat in which it is advantageous. The chances of its becoming lost or submerged are immensely less than with sexual reproduction. . . . Hence it is easy to conceive (1) that in the economy of nature a largely parthenogenetic, or a wholly parthenogenetic, thelytokous race may have a biological (reproductive) advantage under conditions of severe competition ; (2) that from the genetic standpoint it may at most suffer no handicap ; and (3) that from the evolutionary point of view it may actually be more advanced than its bisexual sister races. Hence, whatever controls sex ratios in water-fleas . . . also promotes the evolution of thelytokous races.' It might be added, too, that certain chromosome conditions such as triploidy and irregular sets (aneuploidy) can exist in parthenogenetic but not in sexual forms because, in the latter, abnormal gametes would be the rule.

These 'freedoms' have been bought at a price. Through biparentalism new combinations of characters, that is, of hereditary factors (genes), come into being and afford a richer amount of material upon which natural selection can act; parthenogenesis on the contrary means much stereotypy. In heterogonous forms, leading double lives, these 'freedoms' have not been fully accepted.

Research on artificial parthenogenesis has provided some of the most brilliant and remarkable results ever obtained by experimental biologists. It has greatly increased knowledge of parthenogenesis, has played a leading part in unravelling complicated cytological phenomena, particularly those of fertilisation, and excited the interest of biophysicists and biochemists. But at the same time the types studied have been restricted in number and mostly aquatic, so that the facts gleaned are of limited application to the majority of parthenogenetic animals, which are terrestrial. The number of types reared to maturity is scanty, which makes knowledge of the sex of artificial parthenogones very meagre. The cytological information on artificial parthenogones is patchy and incomplete. No common activating principle has been disclosed from the many agents used nor has the nature of artificial parthenogenesis been elucidated. How far the mechanisms of experimental parthenogenesis and of natural parthenogenesis, and of experimental parthenogenesis and fertilisation, are identical, is unknown ; while of the origin(s) of natural parthenogenesis experimental methods have not as yet revealed anything.

To support the view that parthenogenesis is commoner than is generally supposed enough has been said ; let us glance now at the plethora of researches it presents. For those whose facilities are limited simple types of breeding experiments are possible with common insects such as saw-flies and beetles. High climbing and the exploration of the arid and inhospitable places should yield more cases of geographical parthenogenesis. Those able to exploit more highly specialised breeding and cytological techniques will find reward in rotifers, flatworms, round worms, bristle worms, crustaceans, thrips, book-lice, green-flies, scale insects, beetles, saw-flies, gall flies, gall wasps, ants and molluscs. Combinations of field and laboratory techniques will certainly increase our knowledge of all types of parthenogenesis as well as of sex. The geneticist would also do well to utilise material obtained from parthenogenetic strains derived naturally or artificially. As for the more recondite problems of activation and fertilisation selection operates entirely against the biologist *pur sang* : the future lies with the 'biophysicochemist' unless, that is, the biologist becomes such a one by some kind of metamorphosis.

Finally, a natural question. Of what possible import is parthenogenesis to mankind ? Pincus (1950) gives a part-answer : 'Parthenogenetic reproduction in mammals would have the extraordinary advantage of the production of females

only whose gene complement is exclusively that of the mother supplying the eggs. . . . It is interesting to note that the bee produces males by a natural parthenogenesis. If, by superovulation, egg transfer and artificial parthenogenesis, the desirable milk cow could be made to yield a large offspring of females like herself the land flowing with milk and honey will be available. Its economy will be based on parthenogenesis.'

Applying the matter to mankind itself, the biologist, speaking objectively, would say that though some may hold that there have been mammalian and human parthenogones in the past and, for all we know, may walk unsuspected in our midst today, there is no scientific evidence of such. But in view of Pincus's results with animals, there can be imagined, should there occur a great preponderance of women, the development of an ethic which would allow artificial parthenogenesis as a means of satisfying the maternal instinct. Nevertheless, speaking personally, whatever freedom may be claimed for research, I suggest that, as human beings, we need to see that its applications do not outrage man's highest sensibilities and true dignity. I should prefer that the intellectual effort of practitioners in the new human biology—biologists, doctors, economists, social scientists, statesmen and politicians—in considering the problem of human reproduction, learn from nature, and direct their efforts towards removing the causes of ill-breeding and of imbalance in the sex ratio, thus rendering artificial techniques unnecessary and freeing mankind from the impediments to eugenic and euthetic natural marriage.[1]

References

Artom, C. (1931). *Mem. Accad. Italia. Cl. Sci. fis. ecc.*, **2.**

Astaurov, B. (1936). *C. R. (Doklady) Acad. Sci. URSS*, II.

Banta, A. M. (1939). *Studies on the Physiology, Genetics and Evolution of some Cladocera.* Carnegie Inst. of Washington.

Bataillon, E. (1910). *C. R. Ac. des Sc. Paris*, CL.

Buchner, P. (1911). *Archiv. Zellforsch.*, VI.

Buchner, H. (1936). *Zeitschr. f. ind. Abst. u. Vererbgsl.*, Bd. **72** (1).

Davidson, J. (1929). *Ann. Appl. Biology*, XVI.

Dehn, Madeleine von (1944). *Chromosoma*, **3** (3).

—— (1950). *Naturwissenschaften*, **18.**

Delage, Y. (1902). *Arch. Zool. Exper.*, X (3).

Delage, Y., and Goldsmith, M. (1913). *La Parthénogénèse Naturelle et Experimentale.* Paris.

Fankhauser, G. (1938). *Jour. Exp. Zool.*, **79.**

Gavrilov, K. (1939). *Acta Zool. (Stockholm)*, **20** (2/3).

Goldschmidt, R. (1923). *The Mechanism and Physiology of Sex Determination.* London.

Harrison, J. W. Heslop (1933). *Journ. of Genetics*, XXVII.

Harvey, Ethel B. (1936). *Biol. Bull.*, **71.**

Hertwig, Paula (1920). *Archiv. mikr. Anat.*, XCIV.

Hill, Renate (1948). *Chromosoma*, **3** (3).

Hindle, E. (1911). *Archiv. f. Entwicklungsmechanik*, **145.**

Kasansky, W. K. (1935). *Zool. Anz.*, **110.**

Kling, K. G. (1936). *Experimentelle Untersuchungen über die traumatische Parthenogenese bei Amphibieneiern.* Uppsala.

Lepori, N. G. (1950). *Caryologia*, **2** (3).

Loeb, J. (1913). *Artificial Parthenogenesis and Fertilisation.* Chicago.

Matthey, R. (1948). *Rev. Suisse Zool.*, **55** (1).

Monroy, A. & Montalenti, G. (1947). *Biol. Bull.*, **92** (2).

Monroy, A. (1950). *Scientific American*, **183** (6).

Muldal, S. (1952). *Heredity*, **6** (1).

Parmenter, C. L. (1940). *J. Morph.*, **66.**

Peacock, A. D. & Harrison, J. W. Heslop (1926). *Nature*, CXVII.

Pincus, G. (1936). *The Eggs of Mammals.* New York.

—— (1951a). *Scientific American*, **184** (3).

Pincus, G. & Chang, M. C. (1951b). *Physiol. Rev.*, **31** (1).

Reitberger, A. (1940). *Chromosoma*, **1** (4).

Runnström, J. (1949). *Pubb. Staz. Zool. Napoli*, XXI.

Sanderson, Ann R. & Hall, D. W. (1948). *Nature*, 162.

—— (1951). *Evolution*, V (4).

Schrader, F. & Sally Hughes Schrader (1931). *Quart. Rev. Biol.*, **6.**

Seiler, J. von (1949). *Experientia*, V (11).

Siebold, C. T. E. von (1857). *A True Parthenogenesis in Moths and Bees* (Trans. W. S. Dallas). London.

Suomalainen, E. (1940). *Ann. Acad. Scient. Fennicae*, Ser. A, LIV.

Swann, Mildred F., & Mickey, G. H. (1947). *Proc. Louisiana Acad. Sci.*, **10.**

Thibault, Charles. (1948–49). *Annales des Sciences Naturelles*, XI (1).

Tonnoir, A. L. (1920). *Ann. Soc. ent. Belg.*, LXII.

Tyler, A. (1941). *Biol. Rev.*, **16.**

Ulrich, H. (1934). *Rev. Suisse Zool.*, Bd. **41.**

Vandel, A. (1931). *La Parthénogénèse.* Paris.

—— (1937). *Proc. Zool Soc. London*, **107A.**

—— (1940). *Bull. Biol. de la France et de la Belgique*, LXXIV.

White, M. J. D. (1945). *Ann. Rpt. Dept. Genet. Carnegie Inst. Wash. Yr. Bk., 1946–47*, **46.**

Whiting, P. W. (1945). *Quart. Rev. Biology*, **20** (3).

[1] Delivered Sept. 4, 1952.

SOME ASPECTS OF SETTLEMENT IN THE OVERSEAS DOMINIONS

Address by

PROF. R. OGILVIE BUCHANAN

PRESIDENT OF SECTION E

'SETTLEMENT' is a term which, like many others adopted by the geographer from lay usage, has several distinct, though related, meanings and uses. Currently the geographer's use of the term is tending to restrict itself to the material evidences, particularly the buildings, and perhaps even more particularly the dwellings, of the people who live and work in an area. This is a perfectly defensible usage, appropriate especially in the 'older' lands, and much interesting and valuable recent work, including the whole range of urban geography, bears witness to the fruitfulness of the field. In the 'newer' lands, however, until with the passage of time the settlements in this sense acquire some stability of site and form it is the human processes producing these concrete results that rivet the attention, and in these areas the emphasis in lay usage tends to be on the process rather than on the village or town. This is the aspect on which Isaiah Bowman seized in his own 'Pioneer Fringe' and in the series of studies of pioneer settlement that he stimulated, and it is with settlement in this sense that this paper is concerned.

The basic idea of settlement of this kind is the taking possession of unused or little used land by incomers whose aim and expectation is to utilise the resources of the land, to make homes for themselves and their children, and ultimately to develop as full a range of community amenities as their times, their culture and their numbers can provide. Permanent settlement implies permanent resources and conservational use of those resources, and is pre-eminently based on a stable agriculture. This is not to deny the importance of forest and mineral resources. Forests can, if wisely used, support permanent settlement, though the settlement will be thin and the population small in relation to the area used. But commonly their history in new areas has been different: where they grew in good soil, they were removed to provide land for farming; elsewhere their wealth was ravaged and they were commonly abandoned. Minerals are notoriously a wasting asset, and settlement dependent wholly on minerals is inevitably temporary. Such a relatively stable mineral as coal, even in its more massive occurrences, can produce in time such problems as that of south-west Durham, where a not insignificant population has been bereft of its mainstay. In the Union of South Africa cheap labour, cheap power, technological progress, the rise in gold prices, and, finally, the discovery of the new gold fields of the Orange Free State have pushed the economic exhaustion of the gold resources much further into the future. It remains nevertheless one of the two major problems of the Union. Partly then for these reasons, but partly also because the problems of settlement in the overseas dominions of the British Commonwealth have in fact been overwhelmingly the problems of establishing farming settlement, this paper will concern itself essentially with land settlement, the progress of the occupation and working of new agricultural land in these several countries. This is a problem of many facets, economic, political, legal, administrative, all of them set in the areal framework of the region concerned and against the historical background of the people concerned. To put the matter in a different way, the process of settlement in

the pioneering areas and the pioneering stages reflected the attitudes and capacities of the pioneers themselves and of their distant governments and markets as well as the character of the lands they were attempting to subdue, and that process has left its objective imprint on the geography of the settled areas, and indeed, if we may borrow the thesis of Francis Jackson Turner from the United States, in many less tangible ways in the outlook, the history and the character of the peoples of these dominions.

In so far as land settlement is our topic, a short examination of Turner's views may not be inappropriate. Turner, an American historian, sprang into fame in 1893 with his essay entitled 'The Significance of the Frontier in American History.' A short extract will make vivid his concept of the moving frontier :

> 'Stand at the Cumberland Gap and watch the procession of civilisation, marching single file—the buffalo following the trail to the salt springs, the fur trader and hunter, the cattle rancher, the pioneer farmer—and the frontier has passed by. Stand at the South Pass in the Rockies a century later and see the same procession with wider intervals between. The unequal rate of advance compels us to distinguish the frontier into the trader's frontier, the rancher's frontier or the miner's frontier, and the farmer's frontier.'

It has been well said that Turner supplied for the first time a master key to American history. He saw American development as a series of successive stages towards full utilisation of the natural resources of the land of the United States, accompanied by progressive growth of economic and social organisation. In tune with that, and despite the title of his essay, it is the frontiersman rather than the frontier that focuses his attention. It was the frontiersman in Turner's view who contributed so strikingly to the distinctive character of American democracy, to the consolidation of national sentiment as against sectionalism and to the nature of the development of American religious and intellectual life. But the frontiersman is the product of the frontier, and while in human affairs the trader and the miner may (or may not !) impinge on society in much the same way as the pioneer rancher or cultivator, in so far as a true frontier implies continuity, the trader's frontier and the miner's frontier are not frontiers at all, but merely outposts. The rancher's frontier and the farmer's frontier are the true frontiers, and in the United States, as in the Dominions, they were not merely continuous but they marched with one another.

So far as I know, no geographer and only one historian [1] has attempted to test Turner's concept against conditions in any of the British Dominions and no comparative treatment of the Dominions themselves has been undertaken. Yet each of them has shown something of the same kind of progress as the United States through a moving frontier to a virtually static area of occupation and a high degree of stabilisation of use of the resources of the land ; and some appreciation of the differences in frontier conditions from one to another will provide useful aid towards the understanding of some of the present differences in their cultural geography. This, however, is much too vast a subject for detailed consideration in the short space of one hour, and my purpose at the moment is merely to take one or two of the more obvious points of difference and touch on some of their causes and some of their effects.

Probably the first thing that strikes any student of the colonisation of the Dominions is the length of the period over which it was spread. Champlain founded the first permanent French settlement in Canada in 1608 and van Riebeck established the Dutch East India Company's settlement at the Cape in 1652. At the other end of the scale New Zealand was not declared a British possession until 1840 and the same year saw the arrival at Wellington of the first organised band of settlers. Settlement of the Canadian prairies in any real sense did not begin until the last quarter of the nineteenth century and has been largely the work of the twentieth. There is a spread of two and a half centuries, remarkable not so much for the fact that it is two and a half

[1] Alexander in his *Moving Frontiers* briefly applies Turner's ideas to aspects of Australian development.

centuries as for the changes it witnessed. During that time Holland and France both sank from their proud position as world leaders, and their loss to Britain of the Cape and French Canada was an early symptom of their inability to maintain the rôle of active colonising powers. The United States had grown from a series of separate and often quarrelsome colonies, almost confined to the Atlantic coast, and numbering only a few million people, to a mighty unified nation already setting new standards for the world in economic and social organisation. Britain had lost her American colonies and received in exchange a lesson in colonial administration that she was never to forget, even though she was to have for a century afterwards occasional and partial lapses. Almost immediately afterwards she had had to endure the long strain of the twenty years of the French Revolutionary and Napoleonic Wars. Most significant of all, however, she had initiated the Industrial Revolution and ushered in a new age. From that Revolution she was in the later nineteenth century to reap a very rich harvest, but the immediate picture for three decades after Waterloo was one of greater misery for more people than had ever been known before or has been since. The enclosures, the ruin of the domestic industries by the factories, the absence of control of either working conditions in the factories or housing standards for the factory workers, a savage penal code, the twin influences of laissez-faire and the repressive political atmosphere that was an understandable reaction from the French Revolution, the difficulties, at that time without any precedent, of the change over from war to peace—all these combined with a greatly increased rate of population growth and a drastic redistribution of population to produce among the masses of the people a state of economic and social misery and seething political discontent.

This close-up, with its emphasis on the foreground of the conditions of life of great masses of British people, is very relevant to one aspect of overseas migration, but it dims a background of even greater importance. In the long run the significant elements of the Industrial Revolution were the increased productivity of labour armed with power-driven machinery and the specialisation of labour necessitated by the specialisation of machinery. In combination they provided both the incentive to and the means for the extension of active trading into more and more widely flung areas and into an ever widening circle of commodities. This was the background of the new era colonisation of the nineteenth century. No longer was the settler prepared to go forth into the wilderness to cultivate his own food, make his own clothes and implements and lead a self-subsistent life, limited in material consumption and in intellectual outlook by the character of his own small plot set in its own small neighbourhood. True, such an extreme subsistence character of overseas settlements had never in fact fully existed. Even the New Englanders had early commenced trading, especially in fish, and trading in the product of a highly specialised agriculture had been the essential condition of the foundation and the maintenance of the plantations, but the very limited coastal areas over which these settlements could be successful is the proof of the general point. From about the end of the eighteenth century, then, new colonisation was likely to succeed only if it could produce efficiently commodities that were in demand in the industrial areas and if those commodities were physically and financially able to stand the long transport from the point of production to the point of consumption. With the proceeds of such exports the settlement could import both consumption goods and the capital equipment necessary to promote further development of economic production and social amenities. The logic of that position is easy enough to see from the vantage point of the mid-twentieth century. It was not so easy to see, at any rate in the United Kingdom, up to the middle of the nineteenth century, and the actual process of settlement, particularly in Australia and New Zealand was from one point of view a struggle by the settlers to get command of the right kind of land in the right kind of amount and the right kind of location. Clearly the conception of 'right,' whether in kind, in amount or in location, would change as technological advances enabled

new commodities to comply with the conditions or previously sub-marginal lands to become economic producers, but at each stage there were non-geographical obstacles to be surmounted before new settlement could adjust itself satisfactorily to the changed conditions. In the Southern Hemisphere colonies these obstacles bore in general a certain family resemblance to one another, though South Africa had certain ones peculiar to itself, and in all the colonies the exact nature and degree of common influence varied considerably.

At the Cape the tightly monopolistic Dutch East India Company, intent only on its trade with the Indies, established its settlement as a means to providing fuel, fresh water, meat, vegetables and flour to its own ships on their way to and from the east. Its initial policy of having all the production of these goods done by its own paid servants was soon given up, but its determination to be the sole trader was not, and officials and settlers alike were forbidden to undertake private trading. What the Company had in mind was a small, compact settlement, in which working farmers would produce food on holdings of about 120 acres (60 morgen) and sell it to the Company. It was clear that extension of settlement would progressively decrease the control the Company was able to exercise and increase the expenses of administration and defence, while, further, if production were correspondingly increased, there would be little prospect of disposing of it : ships, even on this high road to the Indies, and even with the multiplying of French and British vessels, were normally too few to support much trade. Control from Europe, though hardly such close and strict control, and the attempt to restrict settlement within narrowly drawn boundaries are features that appear again in the nineteenth century in Australia and New Zealand. In all three countries a leading element, if not indeed the chief element, in the history of settlement was to be the deliberate efforts by the people on the spot to free themselves from both kinds of restrictions.

It would seem that despite various commissions sent out from Holland and numerous reports submitted by governors, the Council of Seventeen never formed an adequate appreciation of the local circumstances at the Cape. Limited areas of reasonably good soil and rainfall, the long summer drought, the necessity for irrigation almost within sight of Capetown gave conditions for farming alien to the Dutch, the German, and even some of the French settlers. By a long process of trial and error farmers had to win through to a command of Mediterranean concepts and techniques. Only in vine-growing and the making of wine and brandy did a measure of success come at all quickly, and that owed everything to the expertise of some of the French and German settlers. At the same time relations had to be stabilised with the previous inhabitants. The Bushmen hunters were themselves hunted and driven out. With the Hottentots, however, owners of cattle and sheep, relations were on the whole friendly, trading in animals became established despite official ban, and the Hottentots, progressively detribalised, set the tradition of a menial class. This tradition, accentuated by the importing of slaves and the complete cessation of immigration from 1707 till the arrival of the British 1820 settlers, was to have a profound influence on the history of South Africa. In the initial stages the hunting of Bushmen and trading with the Hottentots, the inheritance provisions of Roman-Dutch law and the cramping effects of the economic and political system combined to tempt the bolder spirits out beyond the borders of settlement. Here is the origin of the Boers and their trekking habits, a feature of development completely at variance with the established policy and the current desires of the Company. Yet it had to be given some recognition, since it could not be stopped, and, grudgingly, provision was made for land to be held on loan from the Company in blocks as large as 3000 morgen (6000 acres). The rent could in fact be seldom collected and as time went by these holdings virtually acquired the status of freehold. Not uncommonly a man might have more than one, and practise a kind of transhumance from one to another. This was especially so on the western Karoos, with their chronic deficiency of both winter rain and summer

rain. These Boers were essentially cattle farmers, eking out their livelihood by hunting and a minimum of tillage. Their cattle they bred, and augmented by trading with the Hottentots and, later, by raiding the Bantu. Their lack of interest in wool, so marked that various consignments of merinos to the colony were allowed to become diluted and finally lost in uncontrolled breeding with the native hairy sheep, is perfectly understandable in the conditions of the late seventeenth and early eighteenth centuries—there was as yet no factory industry to create a demand for wool that could not be satisfied locally in north-west Europe, and no sign that the Company would be any less arbitrary in buying wool than in buying wheat or wine.

Here, then, was a real frontier and a moving frontier, with northwards and eastwards its own forward screen of elephant hunters and Bushmen hunters. The frontiersman himself, born out of dislike of and resistance to distant authority, achieved almost complete independence with his cattle, his hunting and his distance from the closely settled country of the Cape and the seat of authority at Cape Town, but he achieved it at the cost of an almost self-subsistent life that brought its own penalties, in isolation, impatience of any form of external restraint or even guidance, narrowness of interest and outlook, an extreme fundamentalism in religion and a developed tradition of yielding to the lure of far horizons.

A glance at the relief and rainfall maps of South Africa will disclose a good deal about the physical nature of the frontier area up to about the middle of the eighteenth century and suggest its contrast with Australia beyond the Great Dividing Range and, still more, with the United States beyond the Appalachians. The analogy would have been closer if Australian settlement had commenced at Adelaide and worked eastward, though the advantage would still lie with Australia. Briefly it was a country capable of carrying only a thin population and a scattering of animals, and it needed no very high rate of increase of people or animals to point the desirability of hiving off fresh pioneers who would continue the advance into fresh territory. When the better rainfall areas of the Eastern Cape Province were reached about the middle of the eighteenth century and the Boer advance guards were probing across the Great Fish and then the Kei rivers, contact was joined with the advanced wave of the Bantu, another expansionist people, cattle rearers and part-time cultivators, with a much more highly developed social, political and military organisation than the Hottentots and a recent history of successful territorial advance.

So here was joined a conflict greater in magnitude and fraught with much more serious implications for the future than the previous contacts of the Boers with the Bushmen and the Hottentots. It was conflict in the most literal sense of the term, punctuated by a series of nine Kaffir wars, and the frontier in the narrower sense became and long remained an armed frontier in which the gaining of new land depended on successful fighting. The question was complicated, but not essentially altered, by the assumption of control by England in 1806, and by the consequential injection of the English settlers of 1820 into the advanced border of the frontier zone in the Eastern Cape. In passing we may note that the plan of small tillage farms for these settlers was reminiscent of the original Company's settlement at the Cape, and foreshadowed the later settlements at Adelaide, Dunedin and Christchurch—all of them destined to lose their planned character almost from the moment the settlers landed. The basic difficulty, however, lay in the expansion impetus on both sides of the line of contact and in the irreconcilable, and generally misunderstood, concepts of land ownership. Elsewhere only New Zealand, and there only the north and west of the North Island, repeated this phenomenon in any significant way. New Zealand had its Maori wars, won them in the end, largely in the same sort of way in which the Kaffir wars were won, by the irregular forces of the men on the spot, who could and did develop their own local military tactics to meet their own local situation. In each case victory made more land available for white settlement and left a heritage of resentment and distrust on the part of the defeated. But there the parallel ends :

New Zealand was fortunate in the smallness of the numbers of the Maori, in their high potential capacity for assimilating European civilisation and reaching European standards, and in acquiring a degree of understanding of and sympathy for the native people, which, partial and inadequate as it was, was yet a unique feature of nineteenth-century white-native contacts.

The Great Trek of 1836 and 1837, preceded by several reconnaissance expeditions, was the culmination of the already long history of Boer dissatisfaction with distant control and their ever-present urge to the acquisition of more and yet more land. The basic attraction was the vast extent of better land reliably reported on the summer rain plateaux of what are now the Orange Free State, the Transvaal and Natal. There were indeed new elements in that the control was now foreign, that the British Colonial Office had banned free grants of Crown land in 1832, that the abolition of slavery in 1834 had been accompanied by quite inadequate compensation, and that the freedom for the Boer to deal with the Bantu as he wished was being severely curbed. These gave a political complexion previously absent and promoted an ambition to found new and independent states, expressed in the organised movement of large bodies of people under chosen leaders to the distant destinations of the High Veld and Natal. There is nothing parallel to this in the other Dominions, and nothing in their history to match the train of events set in motion by the fact of the Trek and the attitudes of mind that supported it. The immediate effect in the Eastern Cape frontier zone was a drastic reduction of both people and occupied land. Walker has estimated that ' by September (1837) some 20,000 souls, " the flower of the frontiersmen " as D'Urban called them, had crossed the Orange River.' Their going made room for sheep on their deserted lands, and it was in the few years following the Great Trek that the great wool-growing enterprise of South Africa began to develop. The times were now propitious, Australia was showing the way, and within twenty years the value of wool exports was to outstrip that of all others combined.

Let us now turn to a more direct look at Australia. Sydney was founded simply as a penal camp, and doubtless its isolation, much more pronounced than that of the Cape, was regarded as an advantage rather than otherwise. Phillip had at first only convicts and soldiers, while the first grants of land were made to officers and the first actual settlers were freed men. Here, as in the Cape, the government was the only buyer of grain, and prices to the farmers were arbitrarily fixed by the government. The obvious requirements were free settlers and capital, and official land policy varied according to the views of the governor and his success or otherwise in getting the approval of the Colonial Office. Free grants of land—sometimes large, as witness the famous Macarthur estate at Camden or the 8,000 acres of the Blaxlands—outright sales, conditional sales with rebates according to the number of convicts maintained or free labourers brought from England, and leases were all tried. As late as 1821, nearly half a century after Phillip's landing, the total population was still below 40,000, and less than one acre per head was in cultivation. Clearly progress had been slow, and such as there was had been accomplished with much friction. There were two outstanding difficulties: the poverty of the soil overlying the Hawkesbury sandstone, that stretched away to the north and south of Sydney and westwards to the crest of the Blue Mountain plateau ; secondly, the acute shortage of labour and its frequently unsatisfactory character. Virtually unpaid labour by convicts promised small prospects of advancement for immigrant free labourers, and the mere fact of labour shortage was a discouragement to the entry of capital. Contributory, however, and not insignificant was the inadequacy of survey, which failed to keep up with alienation of land and often made it impossible to prove title to land when it had been issued and to obtain legal protection against squatters.

It was in the eighteen-thirties that the two movements that were so vitally to affect settlement in Australia, and later in New Zealand, came into full strength and into full conflict. The first was the triumph in England of the ideas of Edward

Gibbon Wakefield, their conquest of the Colonial Office and their expression in the formation of companies to organise actual settlements. In Australia South Australia (1836) was the only such colony. In New Zealand Wellington, Nelson, Otago and Canterbury were all founded by such companies, and the latter two benefited considerably from their planning, though there as elsewhere the system was to break down. Wakefield saw colonial settlement as a problem of how to get people and capital on to the land in the right proportions. Australia had the land, England had the capital and the surplus labour. The obvious policy was to sell the land at a 'sufficient price,' use the land fund thus created to finance immigration into the colony, and in due time, but not so soon that they left the existing farms without labour, the labourers would be in a position to buy their own land and employ labour in their turn. Implied in all this are the concepts of compact settlements, small to medium holdings and mixed agriculture on the British model. A prime difficulty not appreciated by the Wakefieldian apostles was how to fix on the 'sufficient price.' If it was to be fixed in relation to the needs of the colony, it might be out of tune with the value of the land and so discourage investment by those who had the capital. If it were fixed to represent a fair value of the land, it would have to vary according to the attractiveness of the land, and in any case might be quite inadequate to promote the development of the colony as visualised by the theorists. Its main weakness, however, was that it completely overlooked the competing attractions of land for squatting beyond the official borders of the colony, and this was the rock on which the Wakefield schemes ultimately broke.

In the eighteen-thirties squatting in New South Wales was not new. It had begun with small men, ex-convicts or free men, squatting on Crown land or even on alienated but not demarcated holdings on the coastal side of the Great Divide. The squatters rapidly acquired a most unenviable reputation, and the term 'squatter' there as in England, became a term of opprobrium. By the thirties, however, the picture was very different. Blaxland, Lawson and Wentworth in 1813 had discovered a practicable, though not easy, route across the Blue Mountains to a 'fine grazing country facing the west.' Others had gone north to the Hunter River and beyond, others again to the south, all to be followed by a stream of successors. The official colony in its expanded form, as proclaimed in 1829, was confined to the Nineteen Counties, and comprised the coastal belt from the Manning to Moruya together with the nearer edge of the western slopes, but already, despite official insistence, this defined an obsolete position.

If it was true (and it was) that the Nineteen Counties were themselves a real frontier area in a strictly pioneering stage, it was no less true that the scantily timbered grasslands beyond the divide promised freedom from cramping governmental restrictions and offered the prospect of great rewards to the successful. Here the sheep, the stabilised but still improving Australian merino, was the pioneering animal, Australian wool was beginning to squeeze the Saxon and Silesian fine wools from the expanding West Riding market, and the 'wool rush' was in full swing. In Roberts's words, '. . . every able-bodied man thirsted for the bush and pined to ride in the dust behind masses of smelling sheep and live on an unchanging diet of mutton chops, unleavened damper and "post-and-rail" tea.' [1] As with the Cape Boers of nearly a century and a half before, this expansion was against government policy, wedded as it was to concentration of settlement and sale of land, but Governor Bourke, in 1835, recognising that it could not be stopped, eased the situation by issuing annual licences conferring the right to graze stock over as much land as the squatter could manage. Unlike the Boer expansion, however, it was a commercial movement in a commercial age, it had little difficulty with natives, though much with nature, and it advanced with spectacular rapidity so that by 1860 there was a continuous pastoral frontier area from Bass Strait to the Gulf of Carpentaria. Squatter now meant a stockman of substance, the backbone of the country, and the squatters had gathered to themselves

[1] Roberts, S. H. : *The Squatting Age in Australia*, p. 11.

the political as well as the economic power of the country.

In New Zealand the process of settlement commenced later and the pattern of development showed marked differences in the two islands. It was the North Island that in 1840 received the first band of colonists organised according to Wakefield's ideas. The forest and swamp of most of the island gave no encouragement to the broad-fronted advance of a pastoral frontier, but demanded rather a slow and painful small-scale war of attrition on those formidable natural obstacles, and success, if and when it came, meant the establishment of small-scale farms in a slow advance of close settlement. There were serious deficiencies in the organisation of the Wellington and Taranaki settlements ; and the later Maori Wars, originating mainly out of conflicting conceptions of land ownership, further delayed progress. The South Island on the other hand provided a remarkable parallel to New South Wales. Nelson (1840, an offshoot from Wellington), Dunedin (1848) and Christchurch (1850) were all designed to demonstrate the superiority of the Wakefield system, and all alike, almost from the day of their founding, found their prosperity in the rise of squatting. The grasslands of the eastern side of the South Island, from the Wairau to the Waiau, with only a handful of Maori at a few scattered points, provided similar opportunities to those of New South Wales, though different difficulties. The tussock grass was far from ideal grazing, but it sufficed for the merino. The country it covered was largely mountainous and the slopes were steep. The rivers, fed from the eternal snow, were torrential and when in spate with the spring snow melt were fearful obstacles. Winter cold and snow and flooded rivers rather than sun-scorched, drought-smitten plains were the hazards faced by the New Zealand squatter. His first sheep and his methods of handling them came largely from Australia, though he himself was mainly not Australian.

By 1860, then, this grassland frontier had been occupied, essentially by a scatter of great sheep runs on no more secure a tenure than short-period licences to graze. Here and in Australia the squatters could not rest content with that hazardous status, and set themselves to the task of getting complete control of the land they depastured. Only in Queensland, and there not until 1916, did any of the governments take the extreme step of abolishing alienation of land, though in the end all of them reached the point of giving long leases of land in the less attractive areas, the arid plains of the far interior in Australia and the difficult ' high country ' of Canterbury and north-west Otago. Elsewhere all governments reserved the right to sell land for closer settlement as population increased and demands for land grew. To evade this danger the squatters adopted many ingenious devices, including the buying of strategically placed pieces of land such as river banks or the land around water holes, possession of which rendered the land behind useless to anyone else. Their measure of success was sufficient to ensure their conversion of most of the desirable land into private estates, consecrated to the golden fleece, but capable so long as it was so consecrated of carrying only an almost negligible population. The squatters in fact had escaped too successfully from the dangers of their insecurity of tenure, and had converted security into a status of privilege that sat ill with the youth of the colonies and the aggressive democracy of the rapidly growing landless populations.

This rapid growth of population was very recent, even in Australia. Not until the eighteen-thirties was there any very significant inflow of free settlers and the disappearance of the distaste and even contempt that the penal settlements of Sydney and Van Diemen's Land had inspired. Capitalistic scepticism of the commercial prospects of these antipodean experiments had been slow to go, and their banishment had not been helped by the early tribulations of the great land companies, the Australian Agricultural Company and the Van Diemen's Land Company, that were given their charters in the middle eighteen-twenties. New Zealand settlement did not begin until the Australian tide had begun to flow, and it started free from the handicap of the early Australian reputation. The discovery of gold, however, in the fifties in New South Wales and Victoria, in the

sixties in Otago stimulated an unprecedented inrush of gold-seekers, few of whom made fortunes from gold, but very many of whom remained to contribute a vigorous, independent and unorthodox element to the existing pastoral society. The significance of this at the time lay in the fact that the colonies had recently acquired responsible government, the solving of land problems was now a matter for the local colonial governments and the power of the ballot box could be used to break the 'land monopoly' that stood between the landless men and the good land fit for close settlement in countries that were destined to support millions of people in comfort. The introduction of refrigeration, in 1882, which opened up vistas of success for small men supplying the European larders with perishable foodstuffs from small farms intelligently and carefully managed, provided a strong reinforcement to the increasing pressure for closer settlement.

The problem was not solved at once. In each of the colonies a long series of measures bore witness to the success of the squatters in defending their interests, but also to the gradual advance towards what had become the community ideal of the family farm capable of providing an acceptable standard of living for a hard-working owner. In detail the measures varied from colony to colony and from time to time, but they bore a strong family likeness to one another. Provision was made against the further alienation of Crown land in large blocks to individuals; for repurchase and subdivision of suitable land already alienated; for survey of farms of size appropriate to the reasonable support of a family, the size naturally varying according to soil quality, to rainfall, to relief and elevation, to accessibility; for assistance to the small settler by government loans at low rates of interest, by an early form of hire purchase known as deferred purchase, or by lease with right of purchase. At the same time conditions of personal occupation and of progressive values of improvements were enforced to guard against the danger of mere speculation in Crown land. The gradual success of such measures ensured the conversion of the better areas of the pastoral frontier zones into an agricultural frontier of close settlement capable of supporting an expanding population. In Australia the over-riding general index to relative productivity was found in rainfall, and the gradual decrease of the rainfall inland has permitted a corresponding, though by no means an exactly corresponding, increase in size of holding from the small dairy farm of the New South Wales South coast, for example, through the progressively larger farms of the sheep-wheat belt and the small grazing run to the great run that remains and will remain undivided in the far interior. In the South Island of New Zealand it was not rainfall but relief and elevation that distinguished the more productive from the less productive land, and broadly the smooth lowlands are wholly occupied by farms, typically within the range of 200 to 500 acres, the lower and easier hill country has small runs of several thousand acres, while the great runs necessarily remain in the high country.

In summary Australia and New Zealand had their period of effective settlement at a time when Europe provided an intense demand for commodities, beginning with wool, that they were well fitted to produce. Ample land, scarcity of labour and complete reliance on a distant overseas market encouraged a severity of specialisation that early became a national habit and still distinguishes the farming countryside from that of South Africa with its entirely different tradition. We cannot doubt that if the nineteenth-century colonisation of Australia and New Zealand had been by Chinese instead of by British people, or if the British colonisation had been seventeenth instead of nineteenth century the rural landscapes of these two dominions would have been very different from what they are.

There is little room left for discussion of Canada, and I must content myself with a brief reference to the Prairie Provinces. The Canadian prairies differed from the Southern Hemisphere colonisation areas in almost every respect. Physically they were a great continuous area of smooth surface, admirable soils, grass vegetation, and a climate characterised by the length and severity of the winter. They were in the far interior of a great continent, separated from long

settled eastern Canada by the hundreds of miles of rough Canadian Shield and from the Pacific coast by the great cordilleran belt, but marched on the south by the expanding settlement of the United States. Yet they had long been used in the fur trade by the Hudson's Bay Company, which enjoyed special rights over the whole area, and indeed until the coming of the railways no other possible product could have been profitably transported to a market. Furs were in some sense the Prairie equivalent of wool in Australia, but sheep were not a reasonable venture in the Prairie climate. The difficulties were well illustrated by the ill-fated Selkirk ventures at settling the Red River valley, even apart from the hostility of the North-West Company. Preliminary measures to settlement in any effective way, then, included resumption of the land by the Canadian government from the Hudson's Bay Company and the building of railways. So it was not until the last two decades of the nineteenth century that the process of settling these grasslands acquired any significant momentum, and even then its precise course and nature depended on three other elements, all of them borrowed from the experience of Canada's great southern neighbour. One of these was the breeding of wheat types to suit the difficult climatic conditions, and but for the success achieved in breeding rapidly maturing wheats almost the whole of the Prairies would have remained beyond the limit of wheat growing. The second was the Homestead Acts, which extended into the Canadian prairies the quarter section of 160 acres and the regular rectangular pattern of fences and roads already standardised on the prairies of the United States. It may indeed be queried whether the adoption of 160 acres as the standard size of farm was altogether wise in view of the shorter growing season and the more restricted range of possible products than in the more southerly United States, and the pretty continuous increase in the average size of holding over the years tends to support that doubt. Nevertheless it was 160 acres that was decided upon, and it proved large enough for the immediate application in one-crop wheat-growing of the contemporary large scale methods of American farmers. Many of the farmer settlers were themselves immigrants from American farms, still following the frontier that had retreated beyond the bounds of their own country, and they were able to demonstrate in practice the economies in prairie settlement of the great machines and the massive teams of horses. On her prairies, then, Canada solved at one operation the legal, administrative and production problems that had cost South Africa some two centuries, Australia over a century and New Zealand over half a century to solve. That does not mean that the pattern of either settlement itself or the production on which it lives is fixed for ever. It does mean that change can be gradual and evolutionary, produced by individual adjustments as opportunity offers or necessity decrees, and not requiring a social revolution powered by a national dissatisfaction and operated by a political party victorious at the polls.

The aim of this paper has been to suggest how differences in the course, pattern and results of settlement in the overseas dominions have been influenced not merely by the inherent nature of the areas in which settlement took place, but very vitally by the human circumstances of the particular time at which the settlement took place. It is a very large canvas, the brush strokes have been very wide, and even so have left most of the canvas untouched. Much valuable contributory work has been done by historians, especially economic historians, but the corresponding definitive work by geographers still offers considerable opportunities.[1]

[1] Delivered Sept. 8, 1952.

FULL EMPLOYMENT IN RETROSPECT

Address by

PROF. D. T. JACK, C.B.E.

PRESIDENT OF SECTION F

IT is perhaps appropriate that an opportunity should be taken on such an occasion as this to bring under review the experience of continuous full employment in the post-war period. The justification may lie in the fact that while in pre-war years there was considerable discussion of the possibility of achieving full employment there was also a certain scepticism to be found lurking on the fringes of the debate, fearful lest the attainment of full employment might involve conditions which would prove unacceptable and apprehensive lest the zealots might be found guilty of generating soft illusions and false hopes. In the event, it has seemed as if the goal had been reached more easily than even the zealots had imagined might be possible, and careful observers are now troubled by the consequences of the very ease with which success has been attained. Certainly seven years of full employment have not been free from crisis. Indeed, by previous standards, crisis conditions have recurred with a more than normal frequency, and to those who find it difficult to suppose that there has been an unusual persistence of mere ill-fortune it is natural to inquire whether the difficulties of this period may not be attributable to the policy, or lack of policy, pursued over the same period in matters connected with the level of employment.

In the pre-war debate the central issue was how full employment might be achieved and the examination of that problem was conducted against a background of prolonged and exceptional under-employment. If the persistence of unemployment which was neither structural nor frictional was attributable to a chronic deficiency of aggregate demand, the ways and means of expanding aggregate demand to the appropriate level had to be explored. What might happen once that appropriate level had been reached was not a problem which obtruded itself unduly into the current controversy. The maxim about first things being first may have suggested that a suitable breathing space would be enjoyed before such awkward matters had to be contemplated. The emergence of inflation and the phenomenon of over-full employment seemed remote possibilities to a world still haunted by the memories of deflation. And the Anglo-Saxon world, while it had made its contribution to the sum total of inflationary experience, had not participated directly in the more extreme forms of inflation which had afflicted much of western Europe within the space of living memory.

Even the new experience of war conditions and their inflationary effects did little to alter the character of the debate. With the recollection of what happened after 1918 it was widely supposed that the cessation of hostilities would be followed by a post-war boom of comparatively short duration and that thereafter, if no special precautions were taken, a post-war depression repeating in the main the features of the inter-war depressions would assert itself. Aggregate demand would grow weary and flag. Hence the emphasis, both in the domestic and in the international fields, on measures which were designed either to maintain an aggregate of demand which would otherwise have faltered or to prevent the spread of economic misbehaviour from the country in which it originated to others which were faithfully and innocently following the path of economic rectitude. But the anticipated situation did not arise. The post-war boom showed no signs of

disappearing. Apart from the economic consequences of uneasy peace, several factors combined to keep employment at a high level. In particular there was an increased propensity to consume, natural enough as a post-war phenomenon but reinforced by fiscal measures which were designed to bring about a more egalitarian distribution of national income ; and there was a high rate of capital investment not balanced by an equally high rate of domestic saving. The result was inflation and over-full employment, with the aggregate of demand pushed to a point which involved severe strains on the economy. There would, I think, be a broad measure of agreement among economists on the ground covered by this brief survey. In what follows I propose to consider certain of the implications of the resulting situation.

At the outset it may be pertinent to inquire whether there is any clearly defined point at which the level of employment passes over from a state of fullness to one of distension or over-fullness. The attempt to answer this question in terms of particular percentages of employment or unemployment is arbitrary, however convenient it may be as a working guide. Lord Beveridge (1) set the fashion in this direction in 1944 when he suggested as ' a conservative aim ' that full employment might be taken to mean the equivalent of 97 per cent. of employment. But Professor Ohlin (2) has rightly pointed out that in full employment the percentage of employment is relatively low where labour mobility is low and where great economic disturbances are occurring and relatively high ' in countries with quiet economic conditions, a well-organised labour market and a skilful adaptation of public works to local surpluses of labour.' The search for precision, however, has been maintained. In 1950, the Economic and Social Council of the United Nations recommended that each government should publish ' as soon and as precisely as is practicable the standard by which it defines the meaning of full employment as a continuing objective of policy, such standard being expressed, wherever possible, in terms either of employment percentages or of absolute numbers of unemployed or in ranges of such percentages or numbers ; and thereafter publish such revised standard as may become necessary from time to time.' An analysis of the replies from governments to a United Nations questionnaire was subsequently published. (3) The Canadian Government declared that it had 'not found it practical to adopt for Canada any arithmetical formulation of a full employment standard ' having regard to its rapidly expanding economy and its peculiar susceptibility to sharp seasonal and climatic variations, fluctuations in crop yields and the vagaries of the export trade in primary products. The British reply referred to a standard, expressed as a percentage of unemployment at the seasonal peak, of 3 per cent. of the total number of employees, equivalent to one-quarter of one per cent. if stated in terms of the total civilian labour force when account is taken of employees and self-employed persons. The United States Government did not determine its standard by reference to any ' definite and unalterable percentage of the labour force ' but had adopted a procedure whereby the President in his economic reports to Congress announced year-by-year employment goals which might act as guides for economic policy. The Swedish Government did not attempt to adopt a full employment standard ' and probably will not do so since, at least with regard to Swedish conditions, it is felt that a rigid fixing of such a standard will not do so much toward helping to maintain the present full employment.' The Belgian Government gave a statistical definition of full employment and estimated that this definition would allow for an unemployment percentage of 2·9 of the total civilian labour force or 3·7 of the labour force available for hire. This excursion into the niceties of the decimal point may be set against the disarming simplicity of the reply of the Kremlin that ' there is no unemployment in the Union of Soviet Socialist Republics and full employment of the population is assured.'

The answer must be sought, not in terms of employment totals or percentages, but in terms of the effects on the economy of the measures taken to increase or maintain the level of employment. These measures consist basically in operating on the flow

of money incomes. Whether or not that is accompanied by public works or by planned expenditure is of secondary importance. In a condition of extensive under-employment, increasing the flow of money incomes though charged with inflationary possibilities does not display the usual symptoms of inflation in the early stages. Some of the 'slack' of unemployed resources will be taken up with prices and wages substantially unaffected. Sooner or later, however, the character of the process alters. Employment continues to increase but prices and wages also move upwards and strains and stresses, which increase in prominence as full employment is approached, appear within the economy. In the course of the same process, moreover, the volume of imports increases under the influence of rising home demand while exportable commodities are diverted into the channels of home consumption and the exporting industries experience increasing difficulties in attracting or retaining labour and capital which would enable them to expand their exports.

Of course there may be special factors which either add to or subtract from this simplified account. In our own case we sacrificed, during the war period, large amounts of our overseas investments and the income therefrom. We received gifts and loans from overseas to assist our post-war recovery and we allowed much of that external aid to be transferred to our overseas creditors. And the adverse movement in the terms of trade obliged us to exchange a larger quantity of exports for a given quantity of imports. But these factors affected the setting of the problem rather than the problem itself and should not distract us from the main issues.

Here we may add, though the point is one to which we shall have occasion to return, that where we are concerned with an overstretched economy the problem of relative prices acquires a greater significance than the problem of general prices because it is movements in relative prices which affect the allocation of resources, and it is that allocation of resources which is so liable to become unbalanced when inflationary forces are at work. I would not suggest that in an under-stretched economy relative prices count for nothing, but by comparison with the situation with which we are now concerned their significance is less. Where there are 'pocketfuls of slack' abundantly scattered throughout the economy an increased demand in one direction is not likely to be satisfied only by withdrawing resources from other lines of production. In a general absence of 'slack' it is a very different type of problem which must be resolved.

Two further considerations may be mentioned. (*a*) When a close approach to full employment has been reached, unless there is an increase in capital proportionate to the further increase in employment—requiring, in the absence of external subventions, an increase in the volume of saving—productive resources will be spread more thinly over the widening field of employment so that we get the emergence of bottle-necks and a tendency for production in the aggregate to slow down. If this occurs while the flow of money incomes continues to increase the inflationary 'gap' is widened and the stage is set for further difficulties to come. (*b*) But the upward movement in prices is unlikely to be allowed to proceed unchecked. In an attempt to restrain the price-wage-price spiral, price controls and subsidies are introduced for what are deemed, for some social reason, to be the more important items of production and the inflation becomes suppressed. At this stage the price mechanism is deprived of one of its main functions in the allocation of resources and this occurs precisely at the time when the problem of relative prices has increased in significance. What I might call the 'natural' distortions of the structure of prices which are a concomitant of the ordinary processes of inflation are now supplemented by the artificial distortions which result from the attempt to suppress the effects of these inflationary forces. These two sets of distortions are not necessarily of the same kind, nor do they necessarily operate in opposite directions. And even if they do, there is no reason to suppose that the second set will exercise a compensating or correcting influence over the first. Bottle-necks, shortages and delays are therefore liable to become even more

conspicuous and in spite of a general appearance of overall activity a deterioration both in quantity and in the quality of production may set in. Excesses of aggregate demand no less than deficiencies can lead to a wastage of productive power. Nor is this type of situation eased by an over zealous concern for the claims of equity, especially if, as is probable, the claims of equity are advanced in disregard of the claims of cost and profitability. Subsidised prices for scarce commodities are calculated to stimulate the demands for such goods and so intensify the shortages. They certainly do nothing to ease the pressure on the limited supplies which are available.

The failure to grapple with the inflationary element in the over-full employment situation has been the principal weakness in the direction of economic affairs in recent years, a weakness which is manifested both by the real strains which are developed within the economy and by the recurring exhortations to industry to step up its productive efficiency. It cannot fail to be a source of surprise that a policy which claims to have raised the level of employment to so high a point should find it necessary, if aggregate supply had kept pace with the increase in aggregate demand, to insist on still greater productivity as a means of preventing an industrial recession or of checking inflation. I hope that nothing which is said here will appear to discourage the many efforts which have been made to respond to these appeals but it must be made clear that increases in productive efficiency are always desirable and that they have no particular relevance as a means of preserving full employment. The fact that special exhortations of this kind are thought to be important by the public authorities may be taken as some evidence that aggregate demand has exceeded aggregate supply and that increased production is required as a means of closing the gap. And we do well to remind ourselves that government has at its disposal an alternative method of achieving the result—by reducing aggregate demand to the appropriate level. Indeed, if the appeal to industry is to be effective it imposes upon government an obligation not to frustrate the response of industry by action which will keep the gap as wide open as it was before the appeal was made.

Apart from the internal stresses created by these inflationary forces there is a likelihood of a persisting balance of payments deficiency which indicates a failure to adjust the economy to the international framework within which it works. In these circumstances, if equilibrium is to be restored, some means must be found for contracting the internal demand for imports to the level of export earnings and for that purpose a variety of measures may be employed ranging from import cuts to new or larger budget surpluses and from higher interest rates to currency devaluation. It is not difficult to conceive of circumstances in which any one of these measures would be preferable to all the others and something will depend upon whether the balance of payments problem is an ephemeral or an enduring one. If it is the latter, as in the situation which we are now considering, the appropriate method may have to be drastic and comprehensive and it must be one which does more than suppress the symptoms. Import cuts, apart from their inflationary nature if taken in isolation, cannot make any strong appeal except as an interim measure pending the introduction of other more effective arrangements. The budget method of restoring equilibrium is slow to operate and cumbersome in its application. Currency devaluation, though on occasion presenting itself as the least objectionable of the available alternatives, is not a suitable instrument for repeated use, and if unsupported by other measures may merely postpone the fulfilment of the crisis to a later date. Higher interest rates, powerful though they may be in their application, may, if relied upon by themselves, create complications in other directions. When crisis conditions exist, however, or when they are developing with speed, it is fatal to refrain from any action because each of the possible actions which may be taken is inconvenient or hurtful in particular quarters. The greater the neglect which has permitted the inflation to develop the more certain it is that the corrective action, when it is taken, will be hurtful.

I come now to two features of the debate

which have received considerable prominence though usually in isolation. One concerns the hopeful search for solutions by the method of international co-operation; the other concerns the tendency for price inflation to be followed by cost inflation within the economy with a consequent apparent shifting of responsibility for the inflationary process, and in particular the hope that salvation may be found through the adoption of a new 'wage policy.'

In both instances the possibility of success depends upon the relevance of the action to the problem to be resolved and, as it happens, in the case of the first, much of the discussion which has taken place has been relevant to a situation which certainly had been anticipated but which has not as yet arisen. It might in fact be said that the approach to the problem of international co-operation had not succeeded in re-orientating itself to the facts of the post-war situation. The point here is one which it is perhaps worth while to develop.

Full employment, according to the now familiar argument, requires the adoption of a particular economic policy at home. But a country with trading relations abroad is exposed to the risk that forces originating outside its territories and over which it has no direct powers of control may inflict significant damage to the particular high level of employment which it was hoped to maintain. It was perhaps true that a diminished foreign demand for the products of the exporting industries might, in terms of aggregates, be offset by an expansion of demand in other directions, but there could be no assurance that the compensating demand could or would be located in a manner which would take up the 'slack' in the exporting industries. Some measure of excuse may be permitted to those who are mildly impatient with the serene dwellers in the realm of economic aggregates and we do well to remind ourselves, in this place, if reminder is necessary, that in certain of these matters, aggregates like patriotism are not enough.[1] On the other hand, a foreign trade may be essential to the most economical deployment of a country's resources and the attainment of the highest living standards so that complete immunity from this vulnerability is not to be contemplated. In these circumstances the problem arises of the possibility of stabilising the external forces which might be responsible for disturbing repercussions on the level of employment. In the post-war years the frustrating recurrence of balance of payments crises in this country has done much to encourage the view, to my mind wrongly, that our difficulties have been in the main of foreign origin.

Under Articles 55 and 56 of the Charter of the United Nations member nations placed themselves under an obligation to take joint and separate action to promote higher standards of living and full employment. If, then, the full employment of one country is placed in jeopardy by reason of forces operating abroad there would be a presumption either that some other member or members were not adequately fulfilling their obligations under the Charter or that there were sources of instability which could not be controlled except by joint action. In recent years, three Reports by committees of experts appointed by the Economic and Social Council of the United Nations have considered means for reducing or regulating this type of instability. The first (4) dealt, *inter alia*, with the problem of stabilising the flow of international trade. The second (5) was concerned with the particular instability of countries which were wholly or mainly suppliers of primary products. And the third (6) was charged with the task of formulating 'alternative practical ways' of dealing with the problem of reducing the international impact of recessions while at the same time having regard 'to the problems of under-developed countries, which are specially vulnerable to fluctuations in

[1] In December 1951, before the recession in the textile industry, the unemployment percentage in Northern Ireland was 8·5 as compared with 1·4 in Great Britain as a whole and 2·8 in Wales and Scotland, the Regions in Great Britain which had the highest rates. In April 1952, after the start of the textile recession, the rate in Northern Ireland was 10 as compared with 2·2 in Great Britain as a whole and 4·8 in the North Western Region where the incidence of the recession was most pronounced. It may, of course, be conceded that Lord Keynes was not unaware of the limitations of an argument which laid its principal stress on economic aggregates. Cf. *General Theory of Employment*, p. 286.

international commodity markets and to related fluctuations in the terms of trade.'

The recommendations contained in the first of these Reports, apart from those which were of a domestic character, fell into two main parts. (*a*) Countries which undertook long-term foreign investment should fix their aggregate foreign lending, both on private and on public account, for five years ahead and should undertake to make good any difference between the actual flow and the anticipated flow, this to be done through the medium of the International Bank for Reconstruction and Development. (*b*) Countries which experienced a reduction in their exports through a falling-off in foreign demand would be entitled to obtain a special allowance of foreign exchange to enable them to meet the deficits on their balances of payments without resorting to cuts in their own imports which would be reflected in reduced demands for the exports of other countries. For this purpose, countries whose imports had fallen off and who were therefore in surplus on their balances of payments on current account would be required to lend back to the deficit countries any increase in their own foreign exchange reserves which resulted from the decline in their imports, this to be done through the medium of the International Monetary Fund.

The measures thus suggested would clearly not be sufficient in themselves since they would be limited to restraining the spread of a recession consequent upon one country's lapse from full employment. Countries whose exports had declined would require to take supplementary measures either to maintain employment in their exporting industries or to absorb in other employments the redundant labour attached to these industries. But that task, though formidable, would be less intractable than the corresponding task where the initial lapse from full employment was allowed to transmit itself cumulatively through a dwindling volume of international trade. Nevertheless the proposals were not regarded as practical by the governments which were represented on the Economic and Social Council. The Report of the third committee—that of the second need not be dealt with here in detail—which was given the task of presenting alternative measures, made three proposals : (*a*) international commodity agreements designed to stabilise particular world commodity markets ; (*b*) an increase in the resources of the International Bank for Reconstruction and Development to enable that institution to expand the flow of its lending in the event of a recession ; and (*c*) increased facilities by the International Monetary Fund to members confronted with temporary difficulties in their balances of payments.

I cite these Reports because they appear to reflect much of the contemporary thinking on the international aspects of full employment policies. Criticism of the Reports has been confined in the main to the practicability of the measures proposed. But there are three comments of a more general character which suggest themselves. First, there is the same disconcerting pre-occupation with economic aggregates which has already been mentioned and which may obscure other quite basic elements in the problem. Foreign lending is regarded primarily as a source of employment. The worthwhileness of the investment in terms of profitability is of secondary importance. As I write these lines there comes to me Professor Frankel's refreshing essay on *Some Conceptual Aspects of International Economic Development of Underdeveloped Countries*. Professor Frankel remarks with justice that ' the discussion about the amount of capital which should be made available by the United States for underdeveloped countries has frequently been based not on the opportunities actually available for its economic *use* but on alleged " calculations " as to how much the United States economy *must* " invest " abroad (in the aggregate!) to " maintain " full employment, export industries, etc., at home.' Such is one of the consequences of ' the view that economic development must be determined by collective objectives expressed in the symbolism of aggregate statistical abstractions.'

Second, it cannot escape notice that the Reports concentrate on a situation in which a significant lapse from full employment has occurred in one of the trading countries. The main purpose of the proposals is to prevent the spread of

recession. What they do not consider with corresponding thoroughness is the situation which arises where one of the countries by an over-zealous pursuit of full employment releases inflationary forces on a scale which generates over-full employment and finds itself confronted with an acute balance of payments crisis. And yet that is the situation rather than the one to which the experts were asked to address themselves which, in our own case, has arisen on three occasions since the end of the War. It may be conceded that the decision to appoint the first committee was influenced by the symptoms of recession which had developed in the United States during the first half of 1949 and which, it was feared, might have widespread repercussions elsewhere. But the other problem had already presented itself on a scale sufficient to warrant the most careful examination and with the speedy recovery of the American economy the case for its examination was more immediately justified, even if the result of the investigation might be mainly negative. For the situation in which a balance of payments crisis arises from inflation and over-full employment at home the proposals put forward by the experts are irrelevant. In such a situation the countries which find themselves in surplus on their balances of payments are not in that position through their own failure to maintain a high level of employment. They are in surplus on account of excessive importing by some other country which has failed to keep its own inflationary proclivities within proper bounds. In these circumstances it is unlikely that those countries which find themselves in surplus will be prepared to lend back to the deficit country whatever increases have occurred in their foreign exchange reserves. If this were not so, there would be no limit to the extent to which an erring country could live on its neighbours. Similarly, the proposal, in the third of the Reports mentioned, to extend the scope of international commodity agreements for the purpose of stabilising particular commodity markets takes no account of the additional strain which could be imposed on such schemes by countries with recurring balance of payments crises consequent upon their own domestic policies. In the post-war years the temptation to attribute one's own misfortunes to the cupidity or neglect of one's neighbours has been exploited by partisan interests to an unpardonable degree, in disregard both of the historical record and of moral candour. International co-operation in the field of economic affairs has an important and useful function to perform, but we do it no service by charging it with tasks which it cannot accomplish. The balance of payments crisis which has recurred in this country is not one which by its nature is capable of solution wholly or mainly by international action, however much its repercussions may be felt by other countries whose governments have taken a stricter view of their economic responsibilities.

Third, and at the risk of some repetition, it may be observed that the measures proposed have a common link with most of the schemes which have been evolved in connection with employment policy in the last eight years. They have involved plans to cope with the type of problem which was prominent during the inter-war depressions. And they have an inflationary bias in so far as they have been more concerned with the maintenance of demand than with the maintenance of supply. It may be that the occasion will yet arise when projects with this type of bias will deserve more practical consideration but to apply them when precisely the opposite type of bias is appropriate is to make for instability and disharmony rather than the reverse.

I come now to the second of the two features of the debate which I mentioned earlier—the tendency for price inflation to be succeeded by cost inflation, and the desire for a new wage policy. In the Report on *National and International Measures for Full Employment* there is some recognition that the general responsibility which falls upon governments for maintaining the stability of the price level is enhanced by their adoption of full employment policies and that to avoid a wage-price spiral 'it is essential that steps should be taken to ensure that the stimulation of effective demand is carried only to the point required to maintain full employment and not in excess of this

amount' (7). To this impeccable sentiment no exception can be taken. There is also some recognition that 'an inflationary price spiral could occur if labour unions, farmers' groups, trade associations, cartels or business monopolies pressed for increased earnings in such a way as to produce a continuous upward pressure on prices' (8). But it is not particularly helpful to read under the rubric of 'recommendations' that 'if there is evidence of a continuous general upward pressure of money wages exceeding substantially the rate of increase of productivity and leading to offsetting price increases, the situation requires such action by the government, jointly with organised labour and employers' associations, as would ensure that any wage increases that may be granted will not result in a general wage inflation. The character of the action to be taken would naturally depend upon the conditions ruling in each country' (9). It is impossible not to contrast the imprecision of this section of the Report with the elaborate confidence with which the recommendations in the field of international action are set out.

The issue here is generally discussed in terms of wage policy. An excess of aggregate demand creates conditions which are favourable to wage increases which push prices up still further. So long as we think of full employment in terms of its popular version 'more jobs than men' we are contemplating a situation which is charged with a high inflationary potential. It is difficult to feel satisfied with the debate which has centred round this issue. Lord Beveridge's solution (10) fell into two parts. (*a*) The central organisations of labour should endeavour to formulate a 'unified wage policy' to ensure 'that the demands of individual unions will be judged with reference to the economic situation as a whole.' This wage policy, we are told, would depend in part upon the behaviour of prices and it should be the business of government to preserve a stable price level. If prices were kept steady, money wages might be allowed to rise in keeping with rising productivity, or at an even greater rate if the share of the national product which goes to labour is to increase. But the emphasis is on a co-ordinated wage policy rather than on a particular wage policy, and we are left with the solace that 'organised labour in Britain has sufficiently demonstrated its citizenship and responsibility to justify the expectation that it will evolve, in its own manner, the machinery by which a better co-ordinated wage policy can be carried through.' (*b*) Under full employment 'wages ought to be determined by reason, in the light of all the facts and with some regard to general equities and not simply by the bargaining power of particular groups of men.' 'This suggests, not that there should be a continuance in peace of the compulsory arbitration which has been accepted in war, but that collective bargains should in general include a clause for arbitration by an agreed arbitrator, in default of agreement between the parties.' [1]

There is, of course, no evidence that organised labour has made or is likely to make any serious effort to formulate a wage policy in the sense in which Lord Beveridge has used that expression. Trade union autonomy in these matters is a highly cherished prerogative which would not be surrendered lightly. However much union leaders may appear as advocates of central economic planning they have no desire to allow this type of planning to be applied to the affairs of their own organisations.[2] In this respect their instincts may speak more wisely than their counsellors. And in any event it is the particular wage policy resulting from the co-ordinated effort to which importance should be attached and not merely the fact that a co-ordinated policy had been evolved. Are we to take it as axiomatic that the co-ordinated wage policy of organised labour could not be other than wise? I find it impossible to subscribe to the simple faith which would answer this question in the affirmative.

In suggesting this scepticism I do not animadvert on the behaviour of trade

[1] For all practical purposes this condition, as regards arbitration, now in fact exists.

[2] It is true, I think, that most exponents of the virtues of central economic planning are reluctant to 'plan wages' and show a tenderness to the trade unions which is noticeably absent from their treatment of other private institutions and activities. An exception to this is to be found in Professor Barbara Wootton's *Freedom under Planning*.

union leaders. We have no ground for supposing that in matters of public responsibility these leaders are significantly different from those who are to be found in other spheres of activity. But their normal responsibilities are towards their members and the entire basis of trade union organisation becomes an anachronism if the main business of trade union leaders in conditions of full employment is to behave as economic supermen in the formulation and execution of a nationally co-ordinated wage policy. Whoever would presume to undertake this task would certainly be least qualified for its performance since the appropriate qualities would include not merely a disinterested regard for the public well-being and the advantage of his own union but also the possession of a degree of knowledge to which no responsible person would lay claim.

The assumption of the proposal is, of course, that a unified wage policy would be more modest in its claims than one which merely emerged out of the diverse pressures exercised separately by the several unions. I cannot with any confidence accept this view and it cannot be supposed that the truth of the proposition is self-evident. And if we look to those instances where a national wage policy has been attempted with or without the concurrence of the trade unions the results are not remarkably encouraging. In Germany, a wage stop was introduced in 1933 when unemployment was high and maintained when unemployment was low, but in the latter period it was subject to much evasion and in any case was accompanied by a drastic curtailment of many of the normal trade union functions. In this country we have had voluntary wage restraints for which some measure of success may be claimed, but if we are to take a long view and contemplate the prospect of permanent boom conditions it is difficult not to suspect that the nebulous principles outlined in successive White Papers (11) will evoke diminishing responses with the passage of time. The mood of the decalogue is hardly a suitable foundation for long-term policy. In the United States the stabilisation efforts of the National War Labor Board from 1942 to 1946 and of the Wage Stabilisation Board since 1950 have been criticised on the ground, among other things, that they were expressed in general rules which lacked flexibility. In Holland, the system of wage control operated by the College of Arbitrators cannot be said to have provided a model for obvious imitation and in the countries of Scandinavia since the lifting of the ban on wage increases it would be difficult to say on what principles the national wage policy has been based.

Despite this discouraging history, the hope still lingers that means may yet be found for making an appropriate wage policy more specific. Running through the British White Papers is the idea that wage movements, other than those which are in response to marked changes in the cost of living, should follow the course of industrial productivity. Lord Beveridge lent his support to this notion in a passage which has already been quoted. Similarly Professor Hansen (12) has suggested, more guardedly perhaps, that ' in some industries it should be possible to include in the collective bargain provision for a systematic schedule for wage increases based on over-all productivity trends ' and that ' this schedule should be subject to periodic review.' In this way, we are told, a rational price policy would become possible. ' If a systematic upward-moving schedule of wage increases were instituted, based on the general gains in productivity for the economy as a whole, then also a continuous pricing programme could be instituted designed to give consumers their share in exceptional cost-reducing improvements through lower prices. In this manner the interests of the general consuming public, together with the interests of employers, could be met by a continuous process of adjustment equitable to all parties concerned.'

In fact, of course, and without reference to undue complications there are in this connection three lines of policy from which to choose. (*a*) The price level might be stabilised and general wage movements allowed to vary directly with productivity. (*b*) The wage level might be stabilised and general price movements allowed to vary inversely with productivity. And (*c*) price and wage levels might be stabilised and hours of work allowed to vary

inversely with productivity.[1] This third possibility may, however, be regarded as a modification of the first and need not be considered separately. If we confine ourselves to the first two we may say that, in the one case, increases (or decreases) in productivity would be reflected by lower (or higher) prices, while in the other they would be reflected by higher (or lower) wages. In the first case there is a primary responsibility upon government to keep the price level stable and perhaps to provide a satisfactory index of general productivity ; in the second there is a greater primary responsibility upon the trade unions to ensure that the wage level shall not be disturbed. On general grounds it is perhaps less difficult to envisage the fulfilment of the first of these responsibilities than the second, however discouraging may have been the experience of recent years, but the margin of confidence is not great.

It is not impossible to contemplate collective agreements which incorporate sliding scales based on an index of productivity, but those who have favoured this type of arrangement have not examined it in detail. Professor Hansen suggested that the device might be introduced ' in some industries ' only, though the index of productivity which he had in mind was one which would reflect ' the general gains in productivity for the economy as a whole.' An alternative possibility is that, for wage-scaling purposes, each industry might have an index of its own productivity. These two possibilities may be considered separately. (*a*) If wages are to be geared to an index of general productivity, assuming that there should be a broad acceptance of the proposition that wages and productivity should be connected by a mechanical link, those industries in which productivity increased by more than the average might be expected to earn more than average profits. It would then fall to the trade unions in those industries to resist the pressure from their members to participate in that greater prosperity by means of even greater wage advances. And on the employers' side, greater prosperity would provide a legitimate ground for an expansion which would require the attraction of a larger labour force, but if they were only allowed to pay higher wages on the basis of the general productivity index for all industry they would suffer a restraint on their ability to attract labour. In so far as relative wage movements influence the occupational and industrial distribution of labour a wage-productivity link of the kind proposed would do damage to the normal working of the wage-mechanism. These considerations may well be conclusive in themselves, quite apart from the unpopularity among wage earners of any arrangement whereby the size of their pay packets are determined by remote principles which cannot be visualised in terms of their own individual efforts. (*b*) If wages are geared not to an over-all productivity index but to an index for the industry in which the workpeople are employed some of the difficulties which have been mentioned become less potent. Trade unions would follow more closely their accustomed rôle and the wage structure would retain a greater measure of flexibility. The objection that industrial boundaries are blurred and that the same occupation often cuts across a number of industries would, on the other hand, be more troublesome. But there remains a more important difficulty. A labour shortage may occur in an industry even though no increase has been recorded in that industry's productivity index. In such a case the industry would be prevented from attracting to itself, by the offer of higher wages, the additional labour which would overcome that shortage. And under conditions of full employment, if inflationary pressures are to be kept in check, the need for a reasonably flexible wage structure as a means of ensuring a proper distribution of labour is even greater than under conditions of less than full employment.

It has already been suggested that in a fully stretched economy the problem of relative prices acquires particular significance. The same is true of the problem of relative wages. To meet this point Professor Lerner has suggested that the

[1] In 1950, the American Federation of Labour invited the Economic and Social Council of the United Nations to place on the agenda for its 10th Session an item concerning the international reduction of working hours as a consequence of rising productivity and this was referred to the I.L.O.

general level of wages might be allowed to rise in proportion to the long-run overall rate of progress in the productivity of labour, that for each industry an 'index of relative attractiveness' should be constructed, and that relative wages should be adjusted by reference to these indices.[1] There is no reason to suppose, however, that the application of this proposal would have either the simplicity or the merit which its theoretical formulation would suggest. Not merely is the number of occupations much larger than is commonly supposed—and I suspect also by Professor Lerner—but their edges do not generally admit of sharp definition.

I am obliged to reject the view that a national or a unified wage policy would provide an answer to the problem which confronts us. On the contrary, there is reason to suppose that such a policy would operate through general rules which, far from making the wage structure more flexible in response to the requirements of a stretched economy, would add to the rigidities which already exist. And if we recognise the assumptions on which the case for a national or unified wage policy is usually based we may reasonably infer that the attempt becomes superfluous if the inflation can be prevented.

The problems to which I have referred in this Address are mainly problems of maladjustment thrown up by the failure to keep a full employment policy within its natural bounds and rendered more intractable by the measures taken to suppress the effects of inflationary forces. A more troublesome problem, however, still remains. The pursuit of a full employment policy carries with it a strong inflationary propensity and it is for consideration whether serious inflation is an essential accompaniment of such a policy. We have the view of Professor Pigou (13) that 'a moderate inflationary tendency, if it enables unemployment to be kept at a low level, would be well worth enduring. It is doubtful, however, whether a moderate inflation would prove sufficient to ensure a really good employment situation.' At least we may say that it is easier to disregard the prudencies and allow the twin evils of over-full employment and inflation to develop than to maintain a 'high and continuous level of employment' in the absence of inflation. But it is the more difficult of these tasks to which our energies should be directed and the experience of recent years should strengthen our resolve. For inflation not merely deals harshly with those whose money incomes are fixed or inflexible; it afflicts the economy in ways which do damage to its efficient performance; and it leaves a legacy of administrative encumbrances not justified in their own cause but only as the hasty and embarrassing response to a situation which need not have been allowed to develop (14). It is no part of the purpose of this Address to examine the monetary and fiscal measures which have been taken in the course of this year to deal with the latest manifestation of crisis conditions. In the long run these particular measures, successful or unsuccessful, will appear as part of the treatment of a short-term situation. But we, and other countries, are committed to the pursuit of full employment as an object of long-term policy, and much remains to be done to ensure that the end may be achieved unencumbered by the instabilities which have been experienced in recent years. We do no service to the public if, in the flush of enthusiasm for the end, we minimise or conceal the perils which attend the adoption of particular measures. Otherwise full employment policies will be taken in Gratiano's fault:

> Thou art too wild, too rude, and bold of voice—
> Parts that become thee happily enough,
> And in such eyes as ours appear not faults;
> But where thou art not known, why, there they show

[1] *Economics of Employment*, pp. 209 ff. The index of relative attractiveness is defined as 'the ratio between the number of people who are qualified and ready to work in the occupation but have not been able to get a job there and the number of people actually employed in the occupation' (p. 214). 'The wage policy then consists of rules indicating that wages should in general rise at a rate equal to the rate of increase in productivity (about 3 per cent. per annum), that wages should rise faster where the *index of relative attractiveness* is sufficiently less than the national average, and that wages should be frozen where the index of relative attractiveness is sufficiently greater than the national average' (p. 219).

Something too liberal. Pray thee, take pain
To allay with some cold drops of modesty
Thy skipping spirit ; lest through thy wild behaviour,
I be misconstru'd in the place I go to,
And lose my hopes.[1]

REFERENCES

(1) *Full Employment in a Free Society*, p. 128. Beveridge. 1944.
(2) *The Problem of Employment Stabilisation*, pp. 5–6. Ohlin, 1950.
(3) *Problems of Unemployment and Inflation.* United Nations, 1951.
(4) *National and International Measures for Full Employment.* Ecosoc. U.N., December 1949.
(5) *Measures for the Economic Development of Under-Developed Countries.* Ecosoc. U.N., May 1951.
(6) *Measures for International Economic Stability.* Ecosoc. U.N., November 1951.
(7) *National and International Measures for Full Employment.* Ecosoc. U.N., December 1949, p. 43.
(8) *Op. cit.*, p. 44.
(9) *Ibid.*, p. 85.
(10) *Full Employment in a Free Society*, pp. 199–200. Beveridge. 1944.
(11) *Price Stabilisation and Industrial Policy.* 1941. Cmd. 6294 ; *Statement on the Economic Consideration affecting relations between Employers and Workers.* 1947. Cmd. 7018 ; and *Statement on Personal Incomes, Costs and Prices.* 1948. Cmd. 7321.
(12) *Economic Policy and Full Employment*, p. 247. Hansen.
(13) *Lapses from Full Employment*, p. 72. 1945.
(14) Cf. Eucken's *This Unsuccessful Age* and Introduction by Professor Jewkes.

[1] Delivered Sept. 5, 1952.

PROGRESS IN AERONAUTICAL SCIENCE AND ENGINEERING

Address by

SIR BEN LOCKSPEISER, K.C.B., F.R.S.

PRESIDENT OF SECTION G

IT is exactly fifty years since the British Association held its last meeting in Belfast. When the President of this Section, Professor John Perry, gave his Address a century was opening that was about to witness more changes in the space of fifty years than had ever occurred in so short a time before. Industrialisation had already come to Britain, and with it the manufacturing power to command the markets of the world. The rising standard of living and the rapidly growing population confirmed the current belief in the inevitability of progress, but nobody foresaw the rapidity of the advance of science and the swift growth and spread of its power throughout the world. Science and engineering have transformed the old social and economic world that John Perry and his audience knew so well, and we are still in the process of adapting ourselves to the new technological civilisation in which we find ourselves.

Nothing so surely puts a stamp on an era as the means of communication, travel and transport. England of the eighteenth century and the stage-coach were all of a piece ; industrialisation and the railways went hand in hand. Now we can fly. It is too soon to estimate the full consequences of so momentous an achievement, but they are bound to be far-reaching. Already we have discovered its consequence in war, for the last decisive military event in our history, the Battle of Britain, was won not, as formerly, on the land or sea, but in the air. In peace, the trend of events is indicated by the striking growth of air traffic. Last year nearly one-third of all the people who crossed the North Atlantic went by air.

Man first took to the air in balloons, but he was at the mercy of all the winds that blew and there was still a year to go, when the British Association last met in this city, before the Wright Brothers flew in a machine that lifted itself by its own power. The flight at Kitty Hawk in U.S.A. was not more than a few hundred yards at a speed of some thirty miles an hour. It is doubtful whether this machine ever climbed more than 10 feet above the ground for it had too much drag, too much weight and too little power to leave much margin for climb. But it was enough. The Wright Brothers succeeded where so many others had failed, because of the relentless scientific attack they made on the problem of flight. As Kelly puts it in his biography of the Wright Brothers ' It is doubtful if the difficulties and full value of the Wrights' scientific researches within their bicycle shop are yet appreciated. The world knows they were the first to build a machine capable of sustained flight and the first actually to fly ; but it is not fully aware of all the tedious, gruelling scientific laboratory work they had to do before flight was possible. Important as was the system of control with which the Wrights' name has been connected, it would not have given them success without their wind-tunnel work which enabled them to design a machine that would lift itself.'

The wind-tunnel was their own invention, in which they measured wind forces and determined centres of pressure on curved surfaces at various angles of attack. They studied the difficult problem of stability, particularly that of lateral stability, and it was their mastery of this problem, and the type of control they invented, which enabled them to overcome

side-slipping and to avoid the tragic fate which overtook earlier experimenters. They were engineers also. They designed and built their aeroplane on the scientific data they had accumulated, and designed and developed their own engine of some 20 h.p., weighing no more than 240 lb.—itself a triumph of engineering. The achievement of the Wright Brothers was in fact the outcome of a rare combination of science, invention and engineering. It came as a climax to a long period of human endeavour and was one of the epoch-making achievements in man's history.

It required no great gift of imagination to visualise the revolution this flight had started. The main principles having been established and successfully demonstrated, progress was rapid. Six years after the flight of the Wright Brothers, Bleriot crossed the Channel in forty minutes, and within sixteen years Alcock and Brown flew the Atlantic, for the first time, in less than seventeen hours. These three machines have changed the course of history.

These three aircraft were probably as good as anybody could have designed and made with the materials, knowledge and techniques available at the time. None of these aircraft, however, could pass muster today. In a space of thirty years or so the aeroplane has evolved from a wooden skeleton, covered in fabric, to a streamlined all metal monoplane built up of hollow shells in aluminium alloy. A glance at the most modern and cleanest of our aircraft, the Comet, in flight, shows the magnitude of the change.

The aeronautical engineer uses the term ' clean ' to signify design excellence, and if we wish to understand what has happened in the air during fifty years we might take this criterion as a starting-point and ask what it implies. A clean design is one of low air resistance, or drag, as it is usually called. This is obviously of the first importance because the work done in level flight between two points is a product of the drag and the distance. The less the drag the smaller the engine thrust required. And since drag varies with air density and therefore decreases with altitude for any given incidence, it pays, over long distances, to fly as high as possible consistent with the load to be carried and the power characteristics of the engine.

Aerodynamic drag is quite a complicated affair, but certain things are obvious from the start. A fixed undercarriage carried in the air stream creates drag, contributes nothing to lift and is a plain liability on the engine. So also are external struts and bracing wires, and, in fact, any parts of the structure that need not be exposed to the air stream. It is the business of the aeronautical engineer to get rid of drag from these causes. Indeed, as the first step to progress he was obliged to do so. He tucked the undercarriage away after take-off. He turned from biplanes to monoplanes and eliminated struts and wires leaving only external wing supports. These, too, disappeared when his increased knowledge of structures, and the production of better materials, brought into being the cantilever wing. He enclosed the cockpit and mounted the engines in nacelles. Now he has buried the engines in the structure, and virtually nothing remains today of drag from unnecessary excrescences.

But a clean design means something more than absence of unnecessary projecting parts. It implies certain standards of excellence in its shape and lines. Streamlined in fact, and the lines we speak of are lines of air flow which, could they be made visible, as they can be experimentally by means of smoke, would be seen, in a perfect shape, to bend smoothly and flow snugly round curved surfaces of the structure, leaving no turbulent wake. Any departure from the ideal aerodynamic shape promotes the breakaway of the air flow from the surface and the formation of large-scale eddies of turbulence. This churning up of the air is sheer waste, and the power to produce it makes an unnecessary call on the engine. How large it was and how great the savings of power which would be possible if profile drag were reduced to its ideal minimum were first pointed out by Melvill Jones in 1929. His lecture on the subject was first regarded as an exercise in idealism, but something approaching what Melvill Jones foretold was accomplished within twenty years.

The other source of drag, known as

induced drag, is bound up fundamentally with lift. Flight is only possible because of shaped bodies called aerofoils whose lift is many times their drag, but however great this ratio, we cannot have lift without induced drag, both of which arise from the circulation of air around the aerofoil. Induced drag can be reduced by using wings of high aspect ratio (the ratio of the span to the chord) but the choice of aspect ratio in design is mainly determined by other reasons.

Profile drag arises from the normal pressures on the surface of a body and the skin friction which occurs in the boundary layer. Ideally, if the flow within that thin layer were laminar over the whole of the surface we should be down to rock bottom. But although great advances have been made we have not attained this ideal state. In laminar flow every particle of air moves tangentially to the surface of the body, and this is not difficult to achieve when the air is first encountered. But in its passage over the wing the boundary layer becomes thicker, and, normally, a transition point is reached when the character of the air flow changes to that of a small-scale turbulence, involving a correspondingly large increase in drag. Figure 1 indicates diagrammatically the various types of flow.

A considerable aerodynamic effort in recent years has been devoted in several countries to achieving as large a percentage of laminar flow as possible, and wing sections have been designed and constructed capable of maintaining laminar flow over 60 per cent. of the chord. The design data in this country have come largely from the brilliant mathematics of Goldstein and the construction from the high-class engineering of the aircraft industry. The conditions required are twofold. Firstly, the pressure conditions along the wing must be favourable—that is, the pressure must fall steadily, and secondly, the surface of the wing must be very smooth and free from waviness. This condition is very exacting as the surface smoothness must be equivalent to that of writing paper, and any waviness of the surface must be restricted to one or two thousandths of an inch, in 2 inches measured in the direction of the chord. Rather more waviness can be tolerated in the direction along the span. Engineering difficulties in producing so true a wing surface and maintaining it in service are obvious enough.

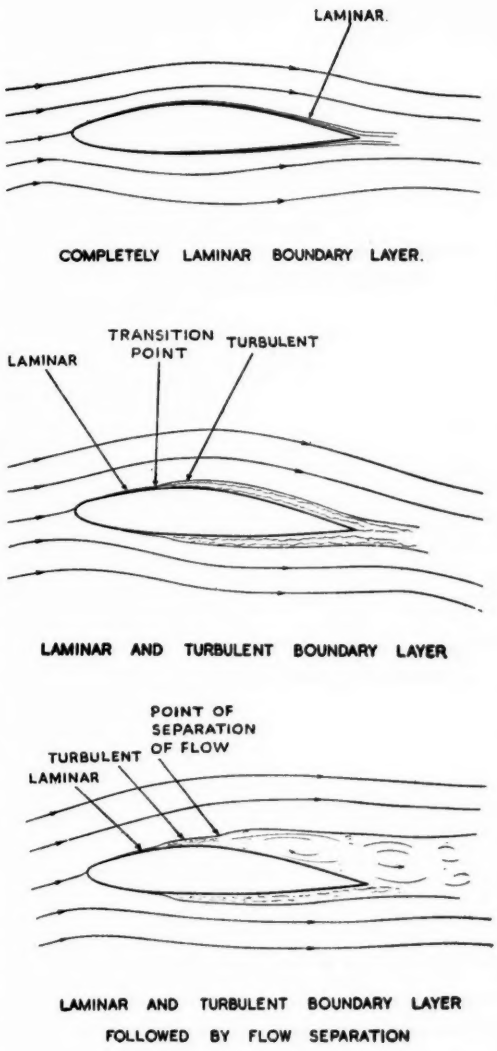


Fig. 1.

It is possible to control the boundary layer and delay the onset of turbulence by sucking part of the layer away through slots or a porous surface. Much effort has been given to the solution of this problem both in the U.S.A. and in this country, particularly by Griffith. Sufficient experimental work has been done to show the possibility, by wing suction, of moving the point of transition from laminar to turbulent flow back to about

75 per cent. of the chord, and if this principle were applied to the best modern design practice we might, at a conservative estimate, expect to reduce the total aircraft drag by about 25 per cent. This is a big factor which, translated in terms of range, would give a gain of 10–15 per cent. for the same take-off weight.

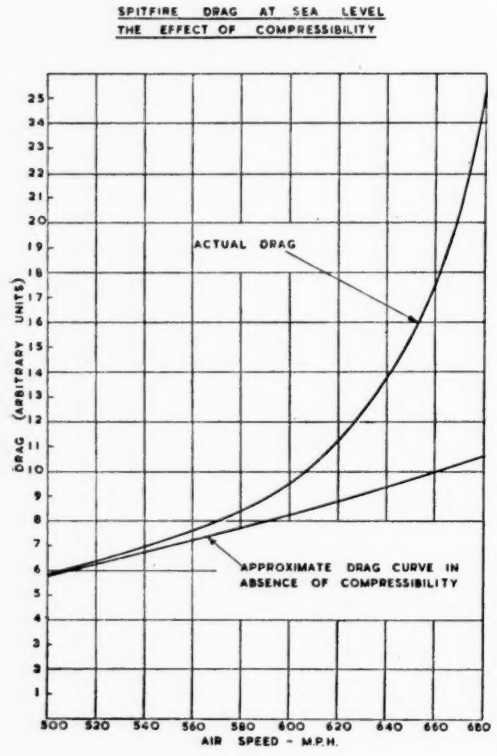

SPITFIRE DRAG AT SEA LEVEL
THE EFFECT OF COMPRESSIBILITY
DRAG (ARBITRARY UNITS)
ACTUAL DRAG
APPROXIMATE DRAG CURVE IN ABSENCE OF COMPRESSIBILITY
AIR SPEED - M.P.H.
500 520 540 560 580 600 620 640 660 680

Fig. 2.

The real obstacle to taking advantage of these scientific and engineering advances arises from the liability of the wing to pick up at low altitudes dust particles and insects which are frequently sufficiently large to cause local breakdown of the laminar flow to the turbulent form. A photograph obtained in flight at the Royal Aircraft Establishment, using an ingenious method, shows up the sharply defined dark triangular patches where local breakdown from laminar to turbulent flow occurs. Each breakdown has been initiated by a minute foreign body at the apex of the triangle. A possible method of minimising this kind of breakdown is to withdraw the turbulent boundary layer through a slot and thus to re-establish laminar flow downstream.

Implicit in the theory underlying the achievement of low drag is the conception of air as an incompressible fluid and for bodies moving at speeds below and not comparable with the speed of sound this conception is substantially true in practice. The velocity of sound in air at sea level is about 760 m.p.h. and the effect of the compressibility begins to make itself sensibly felt above about 500 m.p.h. One effect of compressibility is to produce an increased drag rapidly increasing with rising speed and reaching near the speed of sound a value very much greater than would be encountered under conditions of incompressible flow. Figure 2 derived from data obtained by diving a Spitfire from high altitudes illustrates this.

This is fundamental, for it is impossible to avoid the formation of shock waves which are the source of the largely increased drag, whenever the velocity of the air relative to the surface appreciably exceeds the local speed of sound. This is bound to happen at a forward speed appreciably below the speed of sound in the atmosphere, because the air is accelerated in its flow over a curved body such as a wing. The shock wave occurs at the downstream end of the local region of supersonic flow, and behind it the speed is subsonic. The position of the shock wave depends on the shape of the wing section and its incidence, as well as on the forward speed. The accompanying film, which was taken in a wind-tunnel at the National Physical Laboratory, illustrates the formation of shock waves on a conventional symmetrical aerofoil at high subsonic speeds, and shows the movement of the shock wave back towards the trailing edge as the speed increases at constant incidence. When the shock wave moves back, the low pressure, in the region of supersonic flow ahead of it, extends over a larger fraction of the aerofoil surface, and there is therefore a movement of the centre of pressure with speed and incidence. This causes a considerable change in the stability of the aircraft. The presence of the shock wave may also cause a control surface located at the rear of the aerofoil to become ineffective. At supersonic forward speeds a bow wave is formed, and

the flow again becomes stable with no serious changes in the position of the centre of pressure as the speed or incidence are varied.

The speed of sound in air varies as the square root of the absolute temperature and therefore decreases with altitude. It falls continuously from 760 m.p.h. at sea level until the stratosphere is reached where the speed is 660 m.p.h. A given aircraft speed represents a higher fraction of the speed of sound at altitude than at sea level, and since it is this fraction, rather than the velocity itself, which is the significant factor in compressibility phenomena, it is usual, in dealing with the performance of high-speed aircraft, to speak of the Mach number, i.e. the ratio of the forward velocity to the local velocity of sound.

It is clear that the onset of shock-wave conditions at high subsonic speeds can be delayed by reducing the acceleration produced by curvature, and high speed flight therefore calls for thin wings and slim nacelles and bodies. These increase the critical speed, which can be raised still further by the use of swept-back wings, the effect of which is to reduce the effective thickness-to-chord ratio of the wing in the direction of flight. This great advance in aerodynamic design was first made in Germany, and its effect in delaying the rise of drag is illustrated in Figure 3 (derived from figures given by Sir Harry Garner in his Wilbur Wright lecture, 1952) which compares the drag curves of two hypothetical modern aeroplanes of equal wing areas—one square winged and the other swept back 45°. The diagram illustrates also that the advantage of sweepback disappears at high Mach numbers. A typical modern fighter having swept-back wings and tail surfaces is the Hawker 'Hunter.'

The drag to be overcome at transonic speeds can therefore be considerably reduced by slimming and sweepback, and may be further reduced by more knowledge of the behaviour and control of the boundary layer immediately behind the shock waves, perhaps by suction, to prevent turbulence and breakaway of the flow. The speed of sound was thought at one time to constitute an impassable barrier, partly because of the excessive

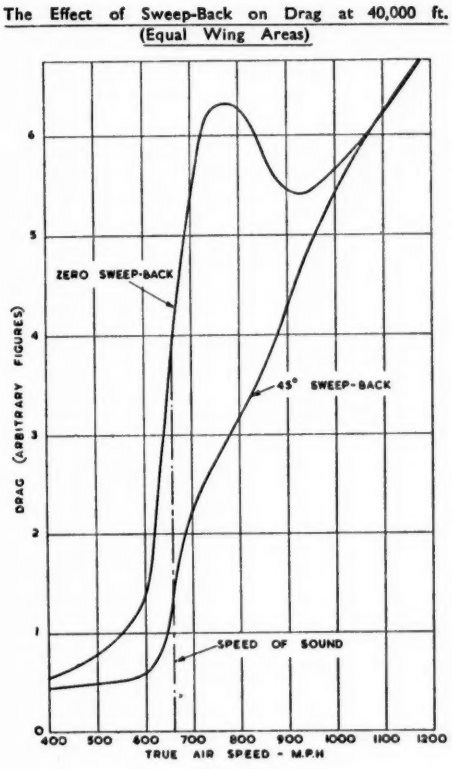
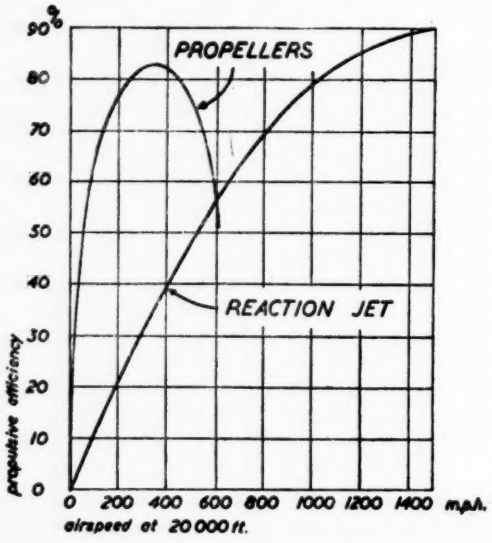


Fig. 3.

drag to be overcome and partly because of the instability and loss of control likely to arise from the changing flow conditions. Both were, of course, very real difficulties, but in the event both have turned out to be less formidable than anticipated. Much of the information which led to surmounting these obstacles came from rocket propelled models and mainly from those experimented on by the National Advisory Committee on Aeronautics in the U.S.A. Experimental aeroplanes both in the U.S.A. and in this country have exceeded the speed of sound.

The air resistance at these high speeds is, however, formidable enough, and the thrust to overcome it has only become possible by the introduction of the gas turbine and propulsion by reaction from a very high-speed jet of gas. Propellers, like aircraft wings, are subject to compressibility drag effects, and the combination of forward and rotational velocities introduces increasing drag conditions over a considerable portion of the blade area after about 400 m.p.h. Up to this point

from comparatively low speeds the propeller is a very efficient means of propulsion, efficiencies of about 85 per cent. being not uncommon, but beyond this, compressibility drag becomes of increasing significance and the efficiency drops sharply. This is shown in Figure 4 which



COMPARATIVE PROPULSIVE EFFICIENCIES

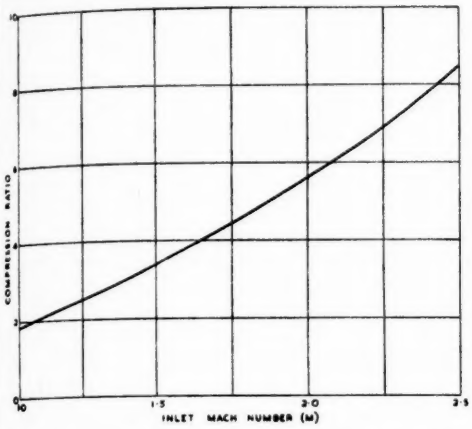
Fig. 4.

contrasts this behaviour with the rise of efficiency of the jet as a means of propulsion.

On this showing the propeller ought to hold its own with the jet up to speeds of at least 600 m.p.h. and, theoretically, even beyond this figure by the application to propellers of the principles of wing design for reducing compressibility effects already referred to. But, in practice, propulsion by jet is ahead of the propeller at lower speeds than this, because of other considerations. The elimination of the propeller eliminates also the disturbance of the air flow over the wing. The turbo-jet engine is much lighter and occupies less space than a corresponding piston engine and propeller, offset though this is by the jet engine's higher specific fuel consumption. Range, of course, plays a determining part in the balance of these factors but it is not the only one. The turbo-jet, particularly with axial flow compressor, can be built with a much smaller frontal area per unit of thrust than any other type of air consuming engine, and it can be made to develop thrusts far in excess of anything hitherto obtainable. These very high thrusts obtained from the jet of the gas turbine, coupled with its small frontal area, are the main reasons why it is so eminently suitable for propulsion at high speeds.

The highest credit attaches to Whittle for his pioneering work on the gas turbine, but it is still in its infancy and its future lies in higher efficiency of compression, higher compression ratios and higher gas temperatures. Advances will come from more knowledge on air flow through rows of aerofoils, involving the same kind of problems of compressible flow as we have already encountered in the flow over a wing. The intake of air in a modern gas turbine engine amounts to about 5 tons a minute, and we are now just as much concerned about the behaviour of air flowing through an aircraft as over it. A heavy price is paid for turbulence and separation wherever they occur. Higher working gas temperatures may become possible with blade cooling, either by air or liquids, and from materials possessing better creep resisting properties at high temperatures. It is anticipated that steady advances in aerodynamics, gas dynamics and metallurgy, together with the development of light-weight rotary heat exchangers, will bring about a progressive reduction in the present rather high specific fuel consumption of the gas turbine.

At supersonic speeds the total head pressure at the intake may become so high as to make a compressor superfluous, and since the turbine is only necessary to drive the compressor, the turbine becomes superfluous also. This leads directly to the simple conception of a shaped tube, called a propulsive duct or ram jet, without any moving parts, as a power unit for supersonic flight. Air is encountered at supersonic velocity and is reduced to subsonic velocity through a shock wave at entry before reaching the combustion chamber. Figure 5 shows compression ratios at various Mach numbers ideally obtainable from this natural compressor. The duct is so designed to allow a suitable reduction of speed through the combustion chamber where fuel is burned under



COMPRESSION ACHIEVED IN A RAM-JET BY A PLANE NORMAL SHOCK WAVE.

Fig. 5.

conditions of constant pressure. The burnt gases are then ejected through the jet pipe. The thermodynamic cycle is thus precisely similar to that of the turbo-jet. Figure 6 shows estimates of variation of overall propulsive efficiency of the ram jet with aircraft velocity, and compares them with the efficiencies of turbo-jets and reciprocating engines with propellers at their optimum cruising altitudes. At low speeds, the ram jet is, of course, very inefficient and some form of boost is required to accelerate the aircraft to a sufficiently high speed at which the ram jet can operate economically.

OVERALL EFFICIENCY OF RAM-JET AS COMPARED WITH CONVENTIONAL METHODS OF PROPULSION.

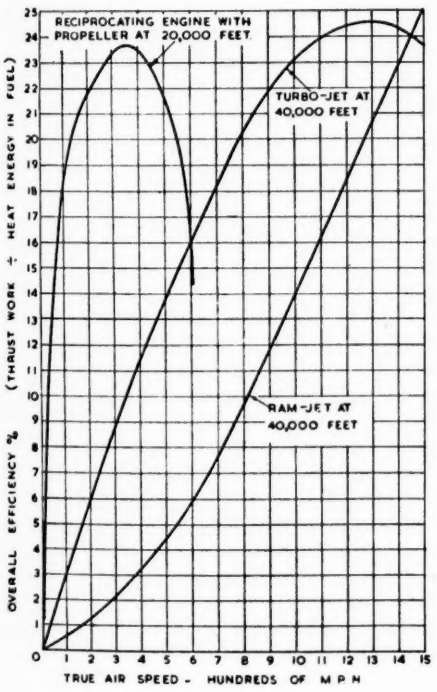

RECIPROCATING ENGINE WITH PROPELLER AT 20,000 FEET
TURBO-JET AT 40,000 FEET
RAM-JET AT 40,000 FEET
OVERALL EFFICIENCY % (THRUST WORK ÷ HEAT ENERGY IN FUEL)
TRUE AIR SPEED - HUNDREDS OF M P H

Fig. 6.

Another type of power unit of limited use in special circumstances in aircraft propulsion is the rocket motor, the characteristic feature of which is that, unlike the power units already described, the oxygen for combustion is not drawn from the surrounding atmosphere but is self-contained within the fuel. The main advantages of the rocket motor are :—

(i) Its comparative simplicity and cheapness.

(ii) The high thrust per unit area of combustion chamber cross section that can be obtained—10,000 to 30,000 lb. per sq. ft. are common, and

(iii) The maintenance of approximately constant thrust with altitude. This is peculiar to the rocket motor and is of particular significance in obtaining high rates of climb at high altitudes.

But of course a heavy price has to be paid for these characteristics. The rocket motor burns fuel extravagantly to provide the maximum possible thrust. Its specific fuel consumption is very high, about 15–20 lb. of fuel being consumed per hour per pound of thrust as compared with a corresponding figure of about 1–3 lb. for the various air consuming jet propulsion systems. This type of propulsion is best suited for special purposes, e.g. for assisted take-off for aircraft, for increasing the performance of interceptor aircraft during the climb and during combat, or for boosting aircraft from the subsonic to the supersonic range.

The rocket motor was developed to a remarkable degree of excellence by Walter in Germany, one of whose motors formed the power unit of the Me. 163 interceptor fighter of the last war. In the first example of an aircraft powered solely by a rocket motor, the main interest in the design (apart from the use of a landing skid) is the large sweepback of the wings introduced eight years ago, and the

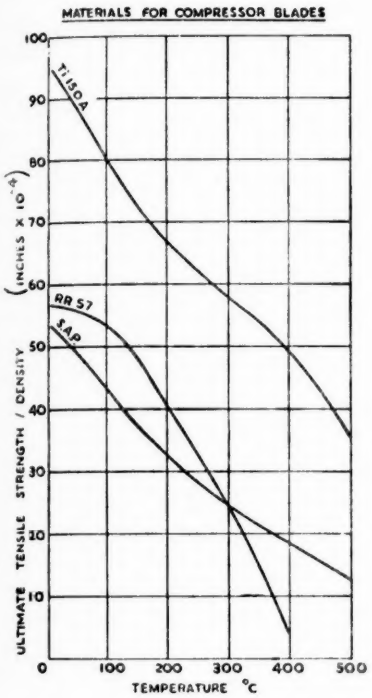
absence of the conventional tail unit. A very similar power unit propelled the German V.2 rocket.

I have already indicated the part the metallurgist might play in the evolution of the gas turbine and many think he holds the key to its immediate future. But there is no doubt of the part he has already played in the evolution of the modern aeroplane. The change from the fabric covered biplane to the superbly finished metal covered monoplane is due primarily to the extraordinary advances made in the properties of light alloys and methods of fabrication. When the Wright Brothers flew, aluminium was little more than a scientific curiosity, a low density weak metal of about 5 tons per sq. in. tensile strength. Modern aircraft aluminium alloys have a tensile strength as high as 40 tons per sq. in. with a density one-third that of steel, and the proof stresses of aluminium alloys have been almost doubled since the time of the first world war—an outstanding achievement providing a striking example of progress made under the spur of necessity.

The development of light alloys which could be rolled into sheet of high strength coincided with important advances in structural science and led to the introduction of the cantilever monoplane which created a major revolution in aircraft engineering. The main structural change was the introduction of stressed skin or monocoque construction. This innovation killed two birds with one stone. It replaced fabric with a more robust material, more capable of resisting deformation and change of aerodynamic shape, and it enabled the surface covering to contribute very substantially to the strength of the structure as a whole. It made the most efficient use of material by placing it as near as possible to the outside surface. The skin now provides nearly all the torsional stiffness, whilst it also contributes to the bending strength of the wing, thus enabling a reduction in the weight of the main spars to be made. The tail surfaces have developed on similar lines, while the fuselage early lent itself to monocoque construction.

The recently developed magnesium alloys, especially those containing zinc and zirconium, have proved of great value in light-weight castings, such as those for the compressor casings of aircraft gas turbines, for they have a strength/weight ratio nearly 20 per cent. greater than that of the best aluminium alloys. Magnesium alloys are also becoming available as forgings, extrusions, and rolled sheet with very promising engineering properties. High specific strength is not their only virtue : they can be machined faster and more easily than any other engineering material, and they have completely displaced aluminium alloys in aircraft wheels. Magnesium alloys now account for about one-third of all light alloy castings supplied to the aircraft and engine industries, and the proportion is growing.

Light alloys based on aluminium and magnesium, suffer, however, a rapid loss of strength with rise of temperature. This has for long been a serious drawback in reciprocating engines where the piston temperature has limited cylinder diameter and power output. In the gas turbine, compression raises the temperature of air leaving a compressor with a compression ratio of 6 : 1 to well over 200° C., and copper-bearing aluminium alloy blades are now giving way to those of titanium alloys containing chromium and iron, whose strength characteristics with temperature, compared with those of aluminium alloys and sintered aluminium powder, are shown in Figures 7 and 8. Titanium alloys which have so far been investigated are twice as strong as aluminium alloys for one and a half times the weight, and further improvements in strength/weight ratio may be achieved. The possibilities ahead in the wider use of titanium alloys, both structurally and for particular use at elevated temperatures, may be great, and justifies a considerable research effort. The cost is at present high—about £5 per lb. as compared with 3*s*. per lb. of aircraft aluminium alloys, but the commercial production of titanium is a recent innovation. The great fall in the price of aluminium from £3 5*s*. per lb. in 1884 to 1*s*. per lb. in 1935 was achieved by the development of new extraction techniques, and it may well be that corresponding advances will be made with titanium. However, the very reactive nature of titanium at the high temperatures which are necessary for its



VARIATION OF ULTIMATE TENSILE STRENGTH WITH TEMPERATURE

Fig. 7.

extraction and working present difficult technological problems, but a very great effort will, without doubt, be devoted to the development of a new metal whose ores are so plentiful and which has such outstanding properties.

The ratio of the elastic modulus to density for present-day titanium alloys is, however, 10 per cent. lower than the corresponding figure for steel and aluminium alloys, and this relative deficiency may, if it persists, prevent full advantage being taken of the high specific strength of titanium in aircraft structures. The development of the modern aeroplane, particularly as regards speed, has introduced a new criterion in design, namely, structural stiffness, which is now a consideration comparable in importance with strength. A reasonable degree of stiffness is essential on practical engineering grounds, but the main concern is with flutter. This is a complicated aeroelastic phenomenon in which, at a particular speed, dependent on the characteristics of the structure, oscillations occur of progressively increasing amplitude, rapidly ending in structural disintegration. Flutter is the skeleton in the cupboard of all high speed aircraft designers, for all structures are liable to flutter in a wind in certain circumstances, and the higher the speed the greater the risk. The best a designer can do is to ensure, as far as he can, that the flutter speed is well above the highest practical flying speed.

The design criteria for achieving this are derived primarily from a masterly mathematical analysis by Frazer and Duncan in 1928 working at the time in the National Physical Laboratory. The mass balancing of the wing about its axis of twist, the close aerodynamic balancing of the control surfaces, the addition of weights to bring their centres of gravity close to the hinge lines, and mechanical irreversibility of the control system—all these measures raise the flutter speed. But increased stiffness, both in flexure and torsion is the principal safeguard as regards the structure itself, and this is the main reason why, as speeds get higher, the need for structural materials of higher ratio of modulus of elasticity to density becomes greater.

The stringent requirements of aircraft engineering and the striking advances made in the structural field have had their influence on other branches of engineering. Flutter is a case in point. Suspension bridges of very long span may also be

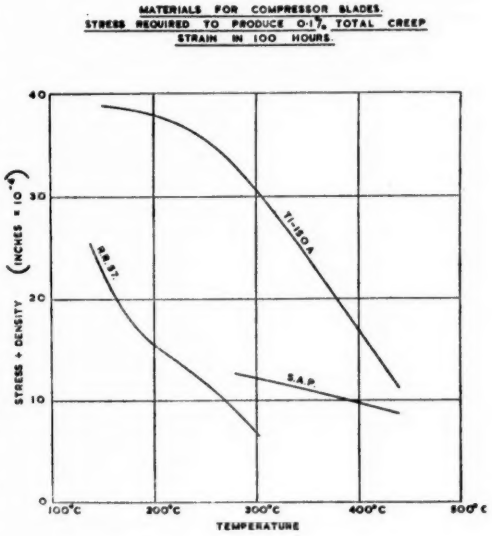


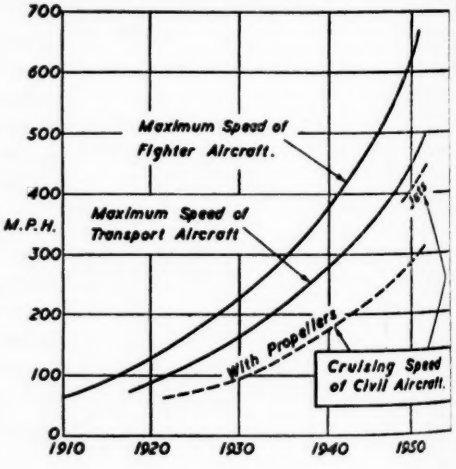
Fig. 8.

liable to flutter at comparatively low wind speeds. The Tacoma suspension bridge in U.S.A. was destroyed in this way in November 1940 in a forty mile an hour wind. This phenomenon of flutter was unsuspected when the bridge was designed, but the designer of long span suspension bridges today has to satisfy himself that the flutter speed is well above the highest wind speed likely to be encountered. This particular problem arose recently in the design of a proposed suspension bridge of 3,240 ft. centre span to cross the Severn Estuary, and has been successfully dealt with, using accurately scaled models of dynamic similarity to full scale in a specially constructed wind-tunnel.

In addition to strength and stiffness at the minimum cost in weight, a third criterion—fatigue—has to be taken seriously into account by the designer in his choice of materials, and this is becoming more and more a matter for anxiety as new alloys of higher tensile strength come into use. The demand for economic air transport urges operators to keep their aircraft in use for as long a period as possible, and flying to schedule often involves flying in bad weather. Considerations of fatigue act as a check on both. Aircraft encounter gusts in the atmosphere and the resultant accelerations impose loads on the structure which may be several times the normal figure. The structure is, of course, stressed accordingly, but although structures can be designed, to a large degree of certainty, to deal with failing loads which need only be applied once, it is quite another matter to deal with repeated loading, particularly when the repetition is irregular, both in time and intensity. We do not know enough, either about the fundamental nature of fatigue or about the frequency and intensity of atmospheric gusts to enable this problem to be treated on an exact basis.

To fill in our knowledge about gusts, many aircraft now carry an accelerometer called 'the counting accelerometer' which virtually counts the gusts encountered in flight and grades them according to severity. A body of information from air routes at various altitudes all over the world is thus being accumulated which enables the designer to assess, with more and more confidence, the kind of repeated load conditions likely to be encountered. In the structural laboratory complete wings are subjected to a fluctuating load broadly representative of flight in bumpy weather, until failure occurs. In addition, joints which are usually the more critical parts of a structure are tested separately in fatigue testing machines. The correlation of the flight records and the test evidence calls for judgment and experience. There is no exact answer and in these circumstances there is often room for disagreement between the scientist and the aircraft operator on the allowable flying life for an aircraft or the critical parts of its structure. Nevil Shute's novel *No Highway* is not all fiction.

I have done little more than indicate some of the scientific and engineering problems that have been encountered in the development of aircraft, and some of the ways in which these problems have been dealt with. As a consequence of these advances in the several fields of science, technology and design, and with a recognition of the vital part played by the skill of pilots, aircraft speeds have increased steadily with the years and, at an increasing rate, from a maximum speed of 70 m.p.h. in 1910 to over 600 m.p.h. in 1950 (Figure 9.) High speed, however, carries certain consequences. High flying speeds involve high rates of



AIRCRAFT SPEEDS

Fig. 9.

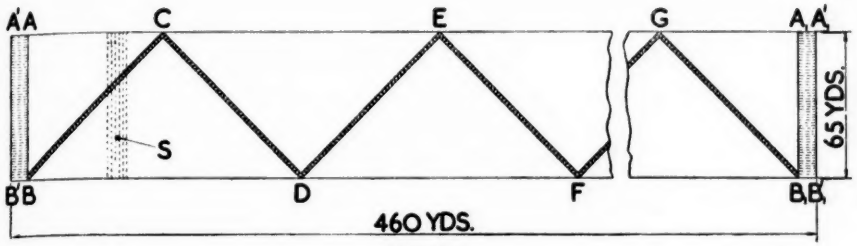


Fig. 10.

fuel consumption, and if we wish to travel fast we have to make it worth while in terms of a reasonable pay load and range. We can get anywhere in the habitable world in hops of about 2,000 miles or so, and as we pay for speed by fuel consumption a fast aircraft at the beginning of a 2,000 mile flight is virtually a flying tanker. The all up weight of the Comet fully laden is about 47 tons of which 46 per cent. is fuel. Further, higher flying speeds involve higher landing and take-off speeds, and fast long-range aircraft are not only inevitably heavy but demand long and expensive runways. Aircraft weight determines the thickness of concrete for a given sub-grade, and take-off speed, for a given acceleration, the length of runway. At the London Airport the longest runway is 3,200 yards, and concrete 12 inches thick is used on a stable gravel base. It is estimated that the total concrete mixed to date for all purposes, including runways, taxi tracks, aprons, road diversions, ducts and drains is about 2 million cubic yards, which is sufficient to build the concrete carriageways of a 20 feet wide 8 inches thick double track motor road from London to Edinburgh. The airport covers nearly 3,000 acres and its cost to date is £10 million. Its estimated cost on completion is of the order of £20 million. If aircraft are to become faster and heavier the main economic problems in aeronautics are, in my view, less in the air than on the ground.

These very heavy runways have led to two interesting developments for decreasing the thickness of construction. The first involves the use of prestressed concrete—a modern technique which has been boldly exploited in many interesting structures. Its application to runways was first tried in 1947 at Orly by the eminent French engineer, Freyssinet. This experimental stretch of runway is shown diagrammatically in Figure 10. It is 60 metres wide and 420 metres long and is built up of precast plain concrete slabs 1 metre square and 16 centimetres thick. Steel wires are threaded through the blocks, transversely to the line of the runway, at 1 metre intervals. These wires are heavily tensioned and anchored to either side of the runway, and so induce transverse compressive stress in the concrete.

The runway is portioned up into large right-angled isosceles triangles (BCD, etc.), and the diagonal joints thus made are closely packed with steel rollers 20 millimetres in diameter inserted vertically between adjacent edges to form linear roller bearings. Longitudinal movement is prevented by deep abutments, one at each end, reaching down some 20 feet or more into the ground. The effect of this is to impose a longitudinal compressive stress of similar magnitude to the transverse stress. The result is a uniform compression of the runway in all directions.

The use of precast slabs, of deep abutments, and of equal prestress in all directions is a costly procedure and experience in England with prestressed roads suggests that considerable economies are possible. A length of prestressed concrete road laid at Crawley by the Development Corporation in 1950, in co-operation with the Road Research Laboratory, uses Freyssinet cables inclined to the centre line to give designed prestressing commensurate with the different transverse and longitudinal requirements. This method avoids the use of precast slabs and of abutments and its cost was similar to that of an orthodox design of reinforced concrete road of twice the thickness.

The second development was suggested

by Williams of the Royal Aircraft Establishment who proposed to lay concrete slabs on water-bags. These bags under the runway, like the inner tubes of motor-car tyres, need only be flexible and impervious ; in neither case are the walls of the inner container required to resist any significant stress. The effect of the arrangement is to convert a concentrated load on the surface of the runway into a uniform pressure on the soil of only 2 to 5 lb./sq. in. (depending on the magnitude of the load), a pressure within the bearing capacity of all but the poorest soils. There are many practical difficulties to be overcome in working out a scheme such as this and it is not at all certain that it would be economical. However, on poor soil it may well be a possible method of construction, and experimental work is now proceeding to explore these possibilities more fully.

The flying boat offers a possible means of escape from large concrete runways and airports. Until comparatively recently the performance of the flying boat was inferior to that of the corresponding land plane, but several advances have now made it a serious competitor. The flying boat always had the advantage of being able to dispense with an undercarriage which, in a modern large aeroplane may account for 6 per cent. of the total structure weight, but this was for a long time more than offset by the excessive size and weight of a hull which had to carry the propellers high above the water to keep them clear of the spray. Jet propulsion has changed all this and a smaller hull, combined with improvements which can now be made in its aerodynamic design, may well tip the balance in favour of a flying boat for all-up weights of 50 tons and above. We are a maritime Commonwealth and suitable waterways for the operation of flying boats could be found in almost all the Commonwealth countries. Flying boats, however, require prepared waterways, and for boats of the size of the *Princess* and *Duchess*, docking and dry-docking facilities. An artist's impression is shown in Figure 11 of a modern flying boat base with facilities for towing the boats to their berths using the Gouge docking scheme. In the absence of detailed studies it is not possible to come to any firm conclusion about costs but many believe that the cost of such a base would be very much less than that of a modern aerodrome, and the advantage of not depriving agriculture of much good land should not be overlooked.

A more radical and indeed revolutionary attack on the runway problem has been proposed by Green and Thomlinson, and worked out in considerable detail at the Royal Aircraft Establishment. The modern

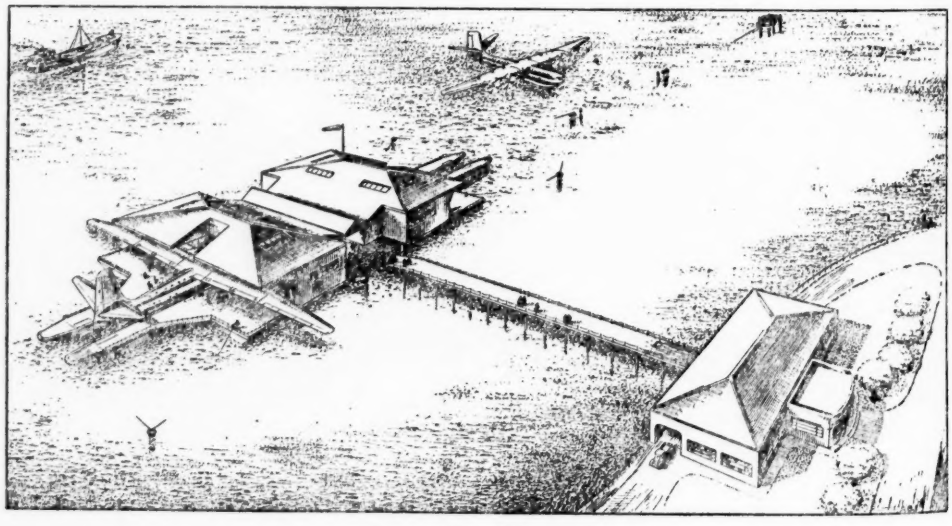


Fig. 11.

concrete runway is an expensive nuisance, however it is constructed, and the heavy and elaborate mechanism of the aircraft undercarriage an unnecessary extravagance in weight and an undesirable handicap in structural design. Green and Thomlinson aimed to get rid of both by catapulting aircraft into the air and 'belly landing' them on a tautly suspended flexible sheet. For landing, the aeroplane is flown level a few feet from the ground, caught on a hook by arrester gear and dropped on the sheet which allows a good landing without bounce to be made.

Some structural reinforcement is required for the fuselage, but this is not expensive in weight, and turns out to be small in comparison with that of the undercarriage. Launching by catapult and the use of arrester gear follows standard aircraft-carrier practice, but the idea of placing on the ground, instead of in the aircraft, the means of absorbing the vertical kinetic energy on landing, is entirely new. Much experimental effort has been devoted to developing this idea and that it is quite feasible is shown by the fact that many landings of a Vampire have been made successfully by this method.

The ability to land on and take off from a small space may become one of the urgent problems of the next twenty years. The helicopter can do this already and that it can do so is the main justification for its existence. From its small operating bases it can travel direct from the heart of city to city. It is not capable of high speeds, as aircraft speeds go, but the Bristol 173 helicopter, which flew for the first time earlier in the year, has been designed to cruise at 120 m.p.h. and this is quite sufficient to give it an advantage in door to door times for journeys such as between London and Paris. British European Airways' previous helicopter services in the U.K. and the coming service between London and Paris are no doubt the beginnings of a more general use of helicopters. But the helicopter may be cast for a more striking rôle. It is at present a very complicated and expensive mechanism, but it could be very much simplified by jet propulsion of the rotor, burning fuel with compressed air at the blade tips. This would involve a higher fuel consumption than at present, but apart from the elimination of the need to balance the rotor torque reaction, the gain in weight saving, simplification and manufacturing costs should be very considerable. A helicopter on these lines has already flown in the U.S.A. With folded rotor wings, as an ingenious German demonstrated before the war, a helicopter could be driven on the road from a private garage like a car to a suitable take-off point. If man is to take to the air in aircraft as he took to the road in motor-cars, this may well be the way he will do it, and long before this century is out a Henry Ford may have mass produced the first of the rotary wing motor-cars.

I confess that these thoughts may have arisen from an uneasy feeling I have that aircraft, the most mobile of all forms of transport, are in some danger of surrendering much of their mobility by becoming tied to expensive and extensive aerodromes. The helicopter, courageously pioneered by Sikorsky in the U.S.A. was for long rather like an unwanted child in the aeronautical family expensively brought up on speed. But this child which has learnt how to operate from a confined space has yet to grow up. We may see, in the not very distant future, the application of folded and faired-in powered rotors to fixed-wing machines, for use only during take off and landing, without prejudice to top speed; and, perhaps, by removing the restraining influence of landing speed on top speed actually make higher top speeds possible. But if this comes about, the primary advantage will lie not in the realm of speed but in the realisation of safer and far more economic landing and take-off conditions. The future in the air may well lie in the conquest of the ground.[1]

[1] Delivered Sept. 5, 1952.

ARCHÆOLOGY, FOLKLORE AND SACRED TRADITION

Address by

PROF. E. O. JAMES

PRESIDENT OF SECTION H

IN these days of intensive and ever-increasing specialisation it becomes progressively more difficult for the President of Section H to find an aspect of his own researches on which to address the Section as a whole which is likely to fall within the compass of the greater part of his audience, drawn as it must needs be from specialists in very diverse fields of anthropological inquiry. This no doubt applies to most of the other Sections, but in the study of so complex an organism as man it is particularly apparent. No single discipline, line of research, or method of approach can be adequate to cover the entire science, or to determine the nature, function and place of the human species in the evolutionary process and in society. Therefore, for the purposes of a Presidential Address it would seem that in concentrating attention upon particular aspects of the subject the wider implications of the underlying principles should be kept in mind, so that the bearing of more specialised studies upon the science as a whole may be recognised. This is not least important at a time when new orientations are being freely discussed.

THE EVOLUTIONARY APPROACH TO THE STUDY OF MAN

Since anthropology came to birth in the latter part of the last century as ' the child of Darwin,' not unnaturally it conformed at first to the tradition in which it was bred, until in the present century, having reached maturity, it has adopted its own lines of development, modes of expression and ways of life. Thus, having begun its career with a very definitely evolutionary outlook, in the changed intellectual climate and social environment of to-day the earlier premisses, presuppositions and bias have undergone considerable modification. Indeed, in some quarters the original approach has been vehemently repudiated. This reaction has been most pronounced in the social and cultural aspects of the science where so many factors are involved calculated to upset any theory of the origin and development of human institutions, organisations and systems as a continuation of evolution in the organic sphere.

In the realms of physical anthropology the Darwinian hypothesis is never likely to be seriously called in question in its main contention since the mammalian origin of the human species is firmly established beyond any possible doubt, however difficult it may be to arrive at agreed conclusions concerning the family tree and the cradle-land of mankind. Although apes and men are only distant cousins, and the common ancestor remains elusive among the aberrant Oligocene and Miocene fossiline specialisations, in anatomical structure man is essentially a primate who, when he attained human status not later than Tertiary times, was in mind, appearance and equipment not much above the simian level. His subsequent development and achievements followed broadly an evolutionary course through the increasing complexity of the Palæolithic cultures amid the vagaries of Pleistocene climate, although it is no longer possible to postulate a straight ascent from the Ape-man of Java, through *Homo Heidelbergensis* and the Neanderthalers, to the Crô-Magnon *sapiens* type, as at one time was supposed.

It is when we pass from the physical to the ethnological and archæological

material that no room for doubt is left that the human situation presents certain peculiarities arising mainly from the greatly increased mental capacity and cerebral development which have characterised man throughout his long and chequered history, and enabled him to control and reflect upon his environment and destiny. From his mammalian ancestry he has derived a common store of instincts, impulses, nervous organisation and mental powers as a continuous evolution. The difference between the human and the animal mind is one of capacity and complexity, but within the human mental endowment lies the potentialities of the species *Homo* as a rational self-conscious creative being. In prehistoric times instinct, emotion and necessity doubtless were more conspicuous than reflective thought but the archæological data reveal that a power to observe record and experiment in a rational, manner was an essential element in the struggle for survival and the mastery of an adverse environment. In adapting the material at hand to his uses in the manufacture of tools and utensils, the specialised techniques he evolved reveal an increasing measure of skill in the art of flint knapping and the production of industries relative to his needs.

Nevertheless, the simple unilineal development of cultures from the Chellean and Acheulean hand-axes to the Mousterian and Upper Palæolithic flake-tools, in the light of our fuller knowledge to-day has had to be abandoned in favour of a much more complicated scheme of intercontinental sequences and traditions. Thus, for example, the flake industry which was first detected at the type-station at Le Moustier in the Dordogne, now appears to have had a long history behind it, and to have been associated with a distinct human stock with Neanderthaloid affinities whose original home was in Asia. Similarly, the Abbevillian-Acheulean core-cultures were characteristic of the ancestors of *Homo sapiens* whose cradle-land may have been in Africa, before the two peoples with their specialised techniques made their way into Europe in the middle of the Pleistocene.

In the light of this evidence, palaeontological, anatomical, ethnological and archæological, it would seem that in a very general way an evolutionary development of man from lower mammals has occurred, conditioned by mental capacity and adjustment to environment, the ability to speak and communicate ideas, to make tools, and eventually to think conceptually. There is no indication, however, in either the biological or the cultural spheres of a *lineal* sequence. While it is now possible to link up one evolutionary stage with another in an interconnected series, the transitions have become infinitely more complex than was formerly assumed, and the resultant types which have survived the ravages of time are the terminal products of many diverging lines of evolution.

Thus, although *Homo Neanderthalensis* disappeared abruptly in Western Europe at the beginning of the Interstadial phase of the Last Glaciation (Würm I, II), when the Mousterian culture came to an end, it is possible that a group of hybrids between palæanthropic and neanthropic types existed in Palestine which played some part in the evolution of *Homo sapiens*. Nevertheless, the typical *Homo Neanderthalensis* appears to have been an aberrant specialisation, or retrogression, which cannot be regarded as an intermediate stage in the development of modern man. Behind it lay a more virile variety of mankind of which the Swanscombe skull is an example, belonging apparently to the Mindel-Riss Interglacial and associated with the Acheulean hand-axe culture. It is to this stock that the direct ancestry of *Homo sapiens* seems to belong, and once it became fully developed and established in the Upper Palæolithic a general mutation in human history may be said to have occurred. Up to this point, although tools often show excellent workmanship, they give no indication of originality, or even change in technique. Local variations occur in different areas of characterisation, sometimes with a very limited conformation to type elsewhere, but the only broad distinction that can be maintained is that between flake-tools and core-tools, connected respectively with two distinct human stocks. The relative consistency and uniformity of the Acheulean as a cultural development over widely separated regions as far apart as

the Somme gravels, Kenya, Stellenbosch and Madras, shows an aversion to innovation, though despite the dominance of an established tradition, intermingling of techniques did occur in particular localities, as in the case of the Clactonian and Acheulean in the Thames Valley of the second Interglacial. Indeed, the distribution and correlation of cultures in different parts of the world suggests that widespread movements of peoples and the diffusion of culture, as well as local evolution, played an important part in Palæolithic development.

The Development and Diffusion of Culture

It was not, however, until the arrival in Europe of the Crô-Magnons, and the groups of *Homo sapiens* allied to them, that the rapid cultural advances were made which characterised the later stages of the last glaciation. In addition to new techniques in the Gravettian, Aurignacian, Solutrean and Magdalenian industries in which bone as well as flint tools were prominent, they introduced plastic human figures in bas relief and animal engravings, together with the mural art which reached its climax in the great Magdalenian polychromes, such as the reindeer at Font de Gaume, the bisons at Altamira and the hunting scene at Lascaux. It is at this point in the development of culture that a sacred tradition becomes manifest, though as early as the Mousterian the ceremonial interments make it reasonably certain that *Homo Neanderthalensis* was *Homo religiosus*, since the ceremonial care with which he laid some of his dead to rest is explicable only on the assumption that he believed in some kind of human survival beyond the grave.

With the advent of the Crò-Magnons however, in addition to the more developed, attempts to restore the body to life by the aid of vitalising agents such as red ochre as the surrogate of blood, perforated shells and disks of stone, found notably at Barma Grande and elsewhere in the Grimaldi series of caverns near Mentone (Verneau, pp. 298 ff.), attention was directed to the magico-religious control of the food supply. Thus, the figures of women with the maternal organs grossly emphasised (often represented in a state of pregnancy) probably are examples of a life-giving symbolism in plastic art which later developed into the fertility cultus of Western Asia and the Ægean, centring in the worship of the Mother Goddess. The distribution of these Palæolithic statuettes in Europe is significant since it supports the view that the Upper Aurignacian culture in which they appeared had a West Asiatic diffusion-centre.

As it came to an end the so-called 'venuses' gave place to the highly developed cave art of the Magdalenians in which the control of the chase was the predominant theme. No one who like myself has visited a very considerable number of these decorated caves throughout the entire areas of their distribution in France and Spain, can be in doubt, I think, that they were primarily sanctuaries of hunting magic and ritual. Except in the case of the Eastern Spanish rock shelters, the inaccessibility of the majority of the galleries in which the paintings occur, and the nature and siting of the designs, are only explicable on the hypothesis that they were executed for the practical purpose of maintaining the food supply by supernatural means, like their counterparts among peoples in a primitive state of culture to-day. The depiction of animals with weapons painted in vulnerable parts of the body, such as the heart, as at Niaux and Le Portel in Ariège, tell their own story.

But the control of the chase was not the only motive. Thus, in the representation in clay of a male bison following a female in the inner chamber at Tuc d'Audoubert, propagation rather than destruction appears to have been the end in view. If the food supply was to be maintained it had to be produced as well as secured. Therefore, as in Central Australia the men of the Arunta tribe resorted to their sacred places in the neighbourhood of Emily Gap to engage in *Intichiuma* ceremonies to promote the fertility of the species on which collectively they depended for their means of subsistence, so the Magdalenians seem to have penetrated to the extreme end of their Pyrenæan sanctuaries for the same purpose—viz., to secure the multiplication of the beasts on which they lived. The rites may have included a sacred dance, if the interlacing heel-marks in

front of the hillock on which the clay models stood can be interpreted as the remains of such a ritual performance (Bégouen, pp. 659 ff.).

That dances of the kind were a feature of the Magdalenian cave ceremonial is shown by the figure of a masked dancer, commonly called 'The Sorcerer,' in the neighbouring cavern known as Les Trois Frères. There we have an actual representation of a shaman engaged in a sacred dance clad apparently in a skin with a long beard, the eyes of an owl, the antlers of a stag, the ears of a wolf, the claws of a lion and the tail of a horse, engraved and outlined in black some ten feet from the ground. In front of it there is an aperture which would make it possible for whatever rites were performed in connexion with the figure to have been seen by those foregathered for the purpose. Moreover, this is not the only instance of representations of sorcerers or divinities in the performance of their functions. Men with animal heads engraved on pebbles and antlers have been recovered from Palæolithic sites at Lourdes, Abri Mège, Marsoulas, Cresswell Crags and La Madeleine. This suggests an established cult in which men and animals were brought together in a mystic relationship in a joint endeavour to conserve and promote the food supply by a ritual technique as the link with the providential sources of life and beneficent abundance.

It seems evident that a great sanctuary like that recently discovered at Lascaux near Montignac on the Vézère was the centre of magico-religious ceremonial for a very long time since almost every phase of Périgordian art is represented from tectiforms to huge polychromes rendered in mauve and purple, as well as in the more familiar black and red pigments, and conventionalised symbolic designs when Magdalenian art began to degenerate into Azilian symbolism and Neolithic abstractions. Here we encounter scenic composition and schematism, which becomes the characteristic feature of the Eastern and South-eastern Spanish paintings of the Cogul-Alperia type, depicting life-like action in which human figures are prominent. Thus, in the hunting story portrayed on the walls at Lascaux it would appear that the death of a hunter who had been killed by a bison and then avenged by a rhinoceros which ripped up the slayer, is recorded, perhaps as the Abbé Breuil conjectures, as a votive painting to the deceased hunter who may have been interred in the vicinity (Breuil, pp. 387 ff.). But whatever may be the interpretation of the galaxy of realistic scenes, symbolic representations, superb friezes, mythical animals and sacred emblems, the cavern must have been a centre of hunting ritual for a succession of ages covering almost the entire period of prehistoric artistic representation. To it recourse was made by generation after generation of hunters to gain control over the food supply and to seek supernatural protection against the hazards of the chase.

With the climatic changes which brought the Pleistocene period to an end about 12,000 to 10,000 B.C., the Franco-Cantabrian naturalistic art disappeared while the schematised tradition continued to flourish and develop in Eastern Spain in the provinces of Tarragona, Castellon, Valencia, Albacete and Alicante, with the survival of a few naturalistic animals and scenes in association with conventionalised human figures. In this region more than 70 shallow rock shelters are still to be seen, and in all probability many more await discovery in the Sierra, containing paintings of this type, of which the best known are the dancing scene at Cogul and the poignant hunting pictures at Alpera. Breuil, in fact, has distinguished no less than six phases in this tradition based on superpositions in the painting, on changes in style and in the species of animals represented. But whether or not the technique arose independently in Eastern Spain contemporary with the Aurignacian and Magdalenian Northern type, or as Professor Pericot of Barcelona now contends, was introduced from Caspian and North African sources, its persistence long after it had reached its zenith, either at the end of the Magdalenian or during the Mesolithic, suggests that it had collected a cultus. Thus, the portrayal of the women dancing round a man (perhaps as a phallic symbol) at Cogul has been painted and re-painted on several occasions, indicating that the shelter was a sanctuary for thousands of years to which

relays of votaries repaired to gain a fresh outpouring of life. Since in some of these schematic paintings human figures are shown wearing masks and horns, the ritual appears to have been much the same as that practised at Trois Frères and other sanctuaries in the Franco-Cantabrian caves.

Nevertheless, as the art became progressively more conventionalised in style doubtless it lost most of its significance in abstraction, so that it was really a spent force when the Neolithic traders and settlers reached Western Europe in the third millennium B.C. Although food-gathering began to give place to food-production in the Near East considerably earlier, only very gradually were the new ways of life, thought and belief diffused and adopted in the outlying regions as a result of the migration of social groups and the movements of individuals. Before a culture that was originally evolved in the Fertile Crescent, the Iranian uplands or the Ægean, could be diffused to Western or Northern Europe it had to undergo drastic modification in the process of adaptation to a very different climatic and human environment. Therefore, a straight development from the Palæolithic to the Neolithic by way of the Mesolithic, and then to the Copper, Bronze and Iron Ages, so congenial to the orderly outlook of a former generation, is a simplification of a highly complex situation which is no longer tenable.

The transition from hunting, fishing, fowling and food-gathering to small-scale cereal-, dairy- and stock-farming was so gradual and intermittent that a very considerable time lag occurred in different localities. Furthermore, only in certain areas, notably in the Ancient East, were the new ways of life widely adopted. Generally husbandry and herding remained for a very considerable time merely adjuncts to the earlier modes of subsistence, and before the cultivation of cereals and the breeding of sheep are possible as a definite economy, the establishment of agriculturists and shepherds in a given region is essential. This presupposes something in the nature of folk movements to explain the gradual spread of Neolithic colonisation from the Eastern Mediterranean, Mesopotamia and the highlands of Iran along the Danube and the Mediterranean littoral to the Balkans, the Ægean, Southern Italy and the adjacent islands, the Apennines and the Iberian peninsula, and thence to the Atlantic coastlands, where the inland winters were tempered by the ocean winds. All this was not accomplished in a day. In fact a thousand years elapsed before the influence of the agricultural civilisations of the Middle East and Southern Europe was felt as far north as the British Isles.

The Megalithic Civilisation

Consequently, it was not until about 2500 B.C. that Neolithic settlers began to arrive from Brittany and the south of France with their cattle, sheep and cereals to take up their abode in their 'causewayed camps' on Windmill Hill near Avebury, and elsewhere on the chalk and limestone uplands. This migration was followed by that of the megalithic builders in search perhaps of metals, markets and new habitats, and bringing with them a highly developed sacred tradition, of which they were the propagators, and which found expression in the massive tombs that characterised their civilisation. The similarity in the structure of these monuments, and in their contents, is indicative of a cultural diffusion, though there is little in their furniture to establish the origin of the distribution. The widely held view that they had passed through an evolutionary history beginning as small simple-chambered dolmens and developing into more complex passage graves, and corbelled vaults before they degenerated into cists, is yet another example of the Darwinian type of thought characteristic of the latter part of the last century which has now had to be abandoned, at any rate in this precise form.

It is true that in Northern Europe the Danish and Swedish megaliths do display a typological sequence along these lines, but only as a local condition. Nowhere outside the Scandinavian region can the simpler tombs be shown to have been earlier in time than the larger tumuli. Thus, in Northern Portugal the free-standing dolmens are for the most part, if not entirely, the remains of concealed tumuli and contain little or no furniture which can be dated earlier than that of the

great southern tombs, though the contents may be of poorer quality. As Professor Daryll Forde has pointed out, they represent ' a provincial degradation typical of peripheral areas,' Furthermore, 'the great galleries and cupolas of Algarve, Huelva, Andalucia and Almeria, the rock-cut tombs of Palmella, and the passage dolmens of central Portugal and western Spain are virtually contemporaneous.' So he concludes that in postulating an evolutionary sequence indicating an Iberian origin for the rough stone tomb ' the classical typology begins at the wrong end ' (Forde, pp. 36 ff., 39). The dolmens or cists of Northern Portugal, Galicia, the Basque provinces of the Pyrenees and Catalonia are in fact the end of the series—mere shadows of the great tombs of the south—rather than the beginning of a process of evolution.

The rarity of metal in the megaliths of Northern Portugal and the adjoining region in Spain, Brittany and North-Western Europe, is no criterion of their antiquity because these regions remained in a Neolithic state of culture when copper and bronze were in common use in the Ancient East and along the Mediterranean littoral (Childe, p. 210). Its absence may suggest that prospecting was not the only or even the chief purpose of the voyagers, though it would seem to have been very often one of their quests. Indeed, from the contents of the graves it is impossible to gauge the nature of their journeyings since no single object or group of objects characterises them. The presence of bell-beaker pottery in association with megaliths from about 2100 B.C. throughout Western and Central Europe from the Iberian peninsula, Brittany and the Rhine to Bohemia, Moravia, and sporadically in Sardinia, Sicily, North Italy and Denmark, shows that the Dinaric Beaker folk had some connexion with the megalithic civilisation.

Being essentially traders who opened up new lines of communication, they diffused the practice of metallurgy which they had adopted in their Iberian area of characterisation, but they were neither metallurgists nor megalith-builders. It was here in Ireland that they acquired very largely the metal weapons and implements they introduced into their settlements, but the Irish ores had been exploited by Iberian smiths before the arrival of the Beaker folk, whose rôle was merely that of fostering trade in metal ware and organising markets for the products of the Irish foundries. They may have encouraged small groups of smiths to settle on the coast of Galloway, and along some of the principal trade routes, but in their own round barrows flint and stone predominated.

Intermarriage between the Beaker people and the earlier megalithic builders may have been responsible for the remarkable development of stone monuments in Britain during this period, though this mode of disposal of the dead not being native to the immigrants (i.e. the Beaker folk) they readily reverted to interment in earth graves, often beneath circular mounds, notwithstanding their association with elaborate megalithic structures like Stanton Drew, Stonehenge, and Avebury, erected for sacred purposes. In Central Europe and the Middle Rhine valley they had intermingled with the Battle-axe warriors of Indo-European origin, and by the time the second wave of these sturdy heterogeneous brachycephals reached the British Isles between 1900 and 1800 B.C. (characterised by the A.C. types of beakers) they seem to have absorbed a considerable proportion of the Battle-axe traditions, together with those of the megalithic civilisation. That this included the adoption of the cultus is shown by the elaborate ritual monuments which they established on the Wiltshire downs. Thus, the Avebury complex and Stonehenge were essentially cult-centres of the megalithic religion from the Early Bronze Age onwards though they had been erected in the first instance by the Beaker folk for the performance of their sacred rites.

The Megalithic Cultus

Precisely what form this ceremonial assumed, and for what purposes it was performed, are not easily determined. That it was rooted and grounded in the cult of the dead is obvious, but this very ancient custom does not alone explain the powerful influence which inspired the movement. If the origin is to be sought in the magnificent galleries, cupolas and

tholoi of Southern Spain and the Ægean rather than in the simple dolmens of the Atlantic coast further to the north and north-west, it is to the sacred enclosure and the nature of the tumulus that we must look for the significance of the cultus. Furthermore, it has to be remembered that the megalithic tradition is a world-wide phenomenon extending from Western Europe to India and the Pacific, and although it cannot be attributed to a single race or group of mankind there are certain common features in the cultus everywhere.

As the genetic relationship of folk tales in Europe and Asia is no longer questioned, it would not be surprising if a similar connexion existed between the various expressions of the cult of the dead and the veneration of heroes in west and east. Thus, the Indian *stupa*, or *tope*, almost certainly originated in the chambered tumuli. When it became a royal tomb it developed greater height, and later (*c.* 250 B.C.) it underwent further elaboration as the central structure of the Buddhist temple, enclosing relics of the Buddha (Fergusson, 1872, pp. 41 f. ; 1910, Vol. I, pp. 62 ff.). Beginning as a circular earthen mound or cairn erected over the mortal remains of a chieftain or holy man surrounded with a balustrade, it was soon covered with plaster, or built of concentric layers of huge bricks and encircled with a wooden railing having gates (*toranas*) on four sides. At an early period it was placed on a circular terrace or plinth for perambulation, and surmounted by a cupola. An umbrella (*chattra*), the Indian emblem of sovereignty, surmounted the tee in the centre of which a stone pillar was embedded, passing through a box-like structure on the top of the dome.

In the plan of the great passage dolmen of Kercado in Brittany, to the north-east of Carnac, some of these features are reproduced. Thus, the tomb is enclosed in a round tumulus and surmounted by a menhir placed over the centre of the interior chamber. The approach to the rectangular chamber is by a megalithic corridor, and surrounding the tumulus is the remains of a stone circle. The mound was composed of fragments of rock faced with ashlar walling, and at the base was a causeway similarly constructed of uncemented stones, flanked by the stone circle (Forde, pp. 70 ff.). Similar structures occur throughout the region (e.g., Mané Lud, Petit Mont, Gavr'inis, Keriavel), and ritual enclosures of this type are familiar at Los Millares, Almizaraque, Belmonte, Tabernas in Spain ; while the Caithness circular cairns and the tholoi of Crete and the Ægean are constructed on the plan of a round hut entered by a low narrow doorway from a small rectangular court surrounded by low dry stone walling. The chamber in these collective graves is usually round (or occasionally rectangular) and contains niches or a porthole for entry, covered with a circular mound or cairn. The corbelled roof is, or was, supported by a stone or wood central pillar, and a forecourt is formed by walling in front of the entrance (Xanthondides). In the Boyne group of Irish chambered cairns, New Grange is surrounded by a stone circle, now reduced to 12 stones but originally numbering probably not less than 35. Around the mound runs a kerb of large flat stones on end, each about 10 ft. long, some of which are decorated with sculpture and geometrical designs. The chamber is approached by a passage 62 ft. in length roofed with capstones and having three cells opening out in a cruciform plan (Coffey, Powell, pp. 243 ff.). When the mound was first opened in 1699 it is said that a menhir lay on the ground in the burial chamber, and another on the summit of the hill (Macallister, p. 127).

In tumuli of this nature the burials within the enclosure consecrated them for the performance of religious rites, the ' fencing ' or stone walling marking them off as sacred ground. As in the agricultural cultus of the Latin farmers in ancient Italy, the sacredness of the boundary stones (*termini*) made them efficacious as a protective device long before they were regarded as the abode of a *numen* or of a *deus Terminus* (Festus, 368 ; Ovid, *Fasti*, ii. 639 ff.), so the supernatural power inherent in the railing kept out all evil and profane influences. Within this sacred enclosure it was the burial or the holy relics that made it a *locus religiosus*, and the greater the person or objects enshrined the more elaborate and venerated the structure, often surrounded with

the meaner graves of subordinates. In fact, once a holy place acquired an established cultus and legend it became either a recognised 'cemetery,' a centre of pilgrimage like a Buddhist *stupa*, or a sanctuary of renown, such as Avebury, Stonehenge, or Bethel, associated with a departed hero, sage or ancestor.

The Legacy of the Megalithic Sacred Tradition

The successive interments so frequently found in barrows and megalithic tumuli show how long they continued to be regarded as a proper place for the burial of the dead. British tombs, for instance, sometimes were enlarged for this purpose (Greenwell, pp. 196, 294, 491), as in the stone cairn of Carn Brea near Land's End, which was used for burial purposes by many peoples of different periods. At the top stood a Christian chapel or hermitage which was destroyed about a century ago. Below were a succession of strata containing human remains ranging from Roman times to the original Bronze Age interment in a stone-lined grave beneath an oval mound (Borlase, pp. v–x). Near by at Porth Curnow, St. Levan, another mound was surmounted by a chapel which contained the funerary urn, very much as the *stupa* was crowned with a cupola, a parasol and a shrine.

At Fimber in the East Riding of Yorkshire a sequence of burials lay beneath the parish church and its two predecessors, extending back through the Anglo-Saxon and Roman periods to the Bronze Age oval barrow which constituted the original site of the holy acre (Mortimer, pp. 189 ff.) Similarly, the little church of Chithurst near Midhurst stands on an artificial mound which rises abruptly out of the valley, and which originally formed a place of assembly for pagan worship, or a pre-Christian cemetery. The present church dates from about 1080 but it is of pre-Conquest design built probably by Saxon masons (Johnston, p. 99 f.). Indeed, it has been argued by A. H. Allcroft that the circular churchyard everywhere has been derived from the round barrow (Allcroft, Vol. I, pp. 25 ff., Vol. II, pp. 305, 310 ff.). This may be an overstatement, but, nevertheless, it is true that churches not infrequently were erected upon circular barrows and conformed to the shape of the site on which they stood.

Similarly, Christian edifices have been built over megalithic monuments, as, for example, at Cangas de Oñis in Asturias, where a passage-dolmen set in a cairn forms the crypt of the small church of Santa Cruz, while about 25 miles from Bilbao is the hermitage of San Miguel de Arrichinaga within the nave of which stands three huge menhirs occupying a space of more than 100 ft. in circumference, and measuring respectively 29 ft., 18 ft. and 14 ft. in height. In the crypt of the magnificent Romanesque church of San Vincente at Avila a tomb is erected above a megalith which is traditionally supposed to represent the rock on which the saint and his sisters suffered martyrdom in 303. Likewise at Padròn, the Roman Iria Flavia, beneath the high altar lies what appears to have been originally a menhir, later inscribed with Roman lettering, and now venerated as the rock to which was fastened the mysterious boat alleged to have brought the body of St. James the Great to the north-west coast of Spain.

The Cult of Santiago de Compostela

Christianity apparently reached the Iberian peninsula before the close of the first century, but little progress was made in the evangelisation of the country until the ninth century. Among the *pagani*, or heath-men, the old religion survived more or less unchanged, and after the Saracen invasion in 710 the Visigoths, who alone had retained their allegiance to the Christian Faith, retired to the mountains and eventually settled in the north-west. There under Pelayo, whom they installed as king of Asturias, they rallied their forces, and in 718 began a series of successful campaigns against the Moors, which in 751 resulted in the recovery of Asturias, Galicia, and Old Castile. To re-establish Christianity in these provinces the cult of St. James, the son of Zebedee, was introduced as a rallying point, so that in the traditional lore of Galicia and Asturias Santiago occupies a position analogous to that of St. Patrick in Ireland.

All the various versions of the cult and its legend agree that after the martyrdom of St. James his body was conveyed in a

ship from Joppa to Iria Flavia (Padròn), where he had engaged in missionary enterprise during his lifetime (cf. Acta Sanctorum, Gelmirez). On arriving at the ancient port on the Galician coast, the mortal remains were placed on what seems to have been a neighbouring megalithic monument consisting of a stone grave and a menhir, venerated locally under the names of *Barca* and *Patronon* ('the "ship" and the "skipper"'). At this point the tradition becomes involved but according to Gelmirez, the first bishop of Compostela, the body was taken to a spot called Libredon, some four miles inland, where it was buried in a tomb constructed of stone from an 'idol' that had been much venerated by the pagans, and over it they built an oratory from the stones that remained of the monument. It now appears from the investigations of Canon Ferreiro that the place selected for the interment was a conical mountain called Monte Illicino where were what are described as 'druidical altars' (Ferreiro). These the disciples of St. James are said to have destroyed and having procured a cart drawn by two bulls they placed the body in it and allowed the animals to conduct it whither they would, as in the Old Testament story of the return of the Ark to Beth-Shemesh. When they reached a region described as Amaia they stopped and buried the body *in arca marmorica*, the local description of a passage dolmen (Leeds, p. 203). The chambers were covered with a mound on which a marble altar was erected.

This accomplished, the tomb is said to have been neglected and forgotten until, about 813, mysterious lights in a copse were seen at night, and sounds of celestial music heard. Upon investigation by the bishop of Iria Flavia a pagan stone 'idol' was found, and beneath it in a cave under a small altar a sarcophagus was revealed bearing the inscription, 'here lies St. James, son of Zebedee and Salome, brother of St. John, whom Herod beheaded in Jerusalem; he came by sea borne by his disciples to Iria Flavia, and from thence on a car drawn by oxen of the Lady Lupa, owner of these states, whose oxen would not pass on any further.' (Lopez, p. 18.)

There can be little doubt I think from the inquiries which I have pursued during the last 25 years in and around Santiago, and in the neighbouring districts on the Portuguese side of the frontier, that Prof. Fleure was correct in his conjecture that the cultus arose out of a tradition of sanctity that in the first instance was associated with the megalithic religion (Fleure, pp. 38, 91; Peake, pp. 208 ff.; Howes, pp. 132 ff.; James, pp. 273). The region abounds in stone circles, dolmens, menhirs and ancient trackways, suggesting that the Galician coast was an important centre from the end of the Neolithic, comparable to Avebury or Carnac. Being near the coast, Santiago would be a convenient meeting place for men from the mountains and men from the sea, as Dr. Howes and Peake have pointed out, just as the fact that it was equidistant from a number of ports in the north-west corner of Spain suggests that, even in prehistoric days, it might have been a junction of roads leading to them, and so, like Canterbury and St. David's, which are similarly situated, have become at an early date a sacred site both for pagan and Christian worship. A common centre promotes not only exchange of goods but also the exchange of ideas, and votaries as well as traders from far and wide doubtless resorted to the Pico Sacro, which probably is to be identified with the Monte Illicino of the legend, with its cave and megalithic remains, and their sacred traditions.

Around the stones known as *Barca* and *Patronon* at Padròn the medieval Santiago cultus seems to have arisen in the first instance (i.e. in the eighth century) to give a Christian significance to the pagan cult. Upon this was superimposed the more elaborate tradition concerning the discovery of the tomb containing the body of St. James in the copse at the place called Libredon, associated with a menhir described as a much venerated stone idol, and the cave containing iron tools. The journey of the body to Compostela by way of this hill is full of megalithic references, and on its arrival at its destination it was buried in a stone tomb in the form of an *arca* or *mamoa* (i.e. dolmen) which became the site of the famous shrine. There eventually the magnificent Romanesque cathedral was erected above the crypt

containing the relics of the apostle, and became the most important centre of pilgrimage in medieval Europe. Thither the faithful made their way, often following the ancient trade routes and trackways along which the Cluniacs established guest houses. It was this religious order which did much to popularise the cult, and it is significant that its popularity was most conspicuous in Brittany, Wales and Ireland where the megalithic sacred tradition was very deeply laid (Hartwell Jones, pp. 244 ff.).

THE MARITIME TRADITION IN PORTUGAL

On their way to Santiago pilgrims frequently made a halt at Orense—the Aurium of the Roman occupation, and a district in which *mamoa* occur, taking its name from the auriferous alluvial deposits in the neighbourhood. The purpose of the 'station' was the veneration of a Byzantine crucifix now preserved in the cathedral and held to have been carved by Nicodemus, who incidentally is one of the heroes of the Arthurian cycle. This sacred object was alleged to have been brought to the Iberian coast from Finistère, an important maritime focal point in the megalithic area in Brittany. Similarly, across the Portuguese frontier another miracle-working crucifix, also attributed to Nicodemus, is said to have reached its present shrine in the church of Bom Jesus de Bouças, Mattoshinos, near Oporto, having come by sea like St. James after a mysterious voyage from Joppa. Under supernatural guidance it floated along the Mediterranean, through the Straits of Gibraltar to the Atlantic, where it turned northwards until finally it landed at Leixões in the parish of Mattosinhos on May 3, 117. The exact spot where it came ashore is marked by a small chapel—*Nosso Senhor de Areia*—situated about half a mile to the north of the harbour. Thence it was carried on the back of a she-ass or mare to Mattosinhos, where the animal stopped, refusing to go another step however severely she was beaten.

From the moment of its arrival countless miracles were performed at the shrine, and notwithstanding the treacherous nature of the coast no wreck, it is claimed, has ever been known at the spot where it came to land at Leixões. Indeed, such was its fame that merchants have insured their ships there, and processions have been held in rough weather, during epidemics, and to stop tempests. It is thought, in fact, to be the oldest shrine in the country, and it still attracts some 30,000 pilgrims annually from all parts of the peninsula.

In this part of Portugal, as in Galicia, megaliths are of simple structure, devoid of corbelling, slab-cists and incised ornament characteristic of the great galleries and cupolas in the south at Algarve and along the Tagus. The workmanship is inferior, the contents are meagre and the distribution is sporadic. As the legends suggest, the movement was from the south making its way originally from the Mediterranean, along the littoral until it came to rest in the natural harbours of north-west Portugal and Spain, while other expansions proceeded to Brittany and the British Isles; the Galician tombs occupying a position midway between those of southern Iberia and Morbihan, diffused by sea around Cape Finistère. The northern Iberian coast was rich in havens like those of Leixões and Padròn within easy reach of tin and gold deposits revealing evidence of ancient mining. Thus, as Santiago was situated at the centre of a mineral region, so the Leça and the Douro offered mining facilities comparable to those of Galicia. Tin-making was extended as far to the south-east as the province of Salamanca, where the mineral occurs under alluvial conditions, as well as in the granite of Valongo, north of the Douro, and near Oporto. Here Cambrian and Silurian strata exist side by side with the metamorphic rock and the granite. Pliny, in fact, described the whole district from the Pyrenees to the Douro as full of mines, tin and gold being especially the products of Galicia and Lusitania (*Nat. Hist.* IV, c. xxxiv; XXXIV, c. xlvii). In the days of Agricola tin was imported from Portugal, and there can be little doubt that the Romans continued a long-established exploitation of mineral wealth on the banks of the Douro and elsewhere in the peninsula, notably in the north-western provinces (Phillips, p. 516). The Romans at Iria Flavia, and their Christian successors at Compostela, almost certainly were following well-trod

paths when they made their way to Santiago, situated on a main route from Coruña vîa Pedròn and Pontevedra to Vigo and the adjacent mineral deposits, in a district populated originally perhaps by voyagers from the south in the Bronze Age.

The 'Citanias' of the Minho

As regards the region south of the Minho in Portugal, the gorge-like *rias* give place to a coastal plain affording easier access to the hinterland if less in the matter of natural harbours. Proceeding inland from the neighbourhood of Leixões, at the mouth of the Leça, there is evidence of important ancient settlements in the Province of Minho around Taipas (renowned for its thermal mineral springs from Roman times onwards) between Guimarães and Braga. Thus, on the spur of the Serra da Cabreira, some 2½ miles above the village of Briteiros, there stand ruins of an ancient Iberian town, commonly called *Citania*, consisting of a series of paved streets flanked on either side by the foundations of circular and rectangular houses, arranged in groups around a court, and with an entrance from the street. One of these houses has been reconstructed by Dr. Martin Sarmiento as a beehive structure with a conical roof, the inside forming a single chamber. Various explanations of their purpose and significance have been offered, among which is that of Professor Altamira. He maintained that since the earlier generations of this people buried under dolmens which were circular when covered, the later folk retained the tradition of circular sepulchres with a domed roof, the middle of which was supported by a pillar of stone or wood, like the indian *stupa*, the entrance corresponding to the passage of the dolmen (Altamira, p. 50). Be this as it may, three dolmens occur below the settlement, suggesting that the site was occupied long before it became a Celtiberian stronghold, just as it continued to be inhabited after the Roman influence dominated the district, probably until and after the establishment of the Visigoths in Portugal. On the summit of the hill stands a little chapel dedicated to São Romão with an inscription, *Aqui foi a antiga capela de S. Romoo* (Here was in ancient times the chapel of São Romão). The foundations of this ancient Christian edifice are still traceable.

According to a document dated 1059 (*Libro de Mumadona*) there was on the hill a hermitage of São Romão belonging to the Benedictine monastery of Guimarães, founded in the tenth century by a Leonese princess, on the site of the ruined ancient Romano-Celtic town. In the ruins of the chapel there is unmistakable evidence of an early Christian occupation of the hill, while the local folklore shows that the buried city was regarded with supernatural dread. Thus, it was thought to be the abode of witches and enchanted Moors who proceeded thence at night to carry evil over the plain beneath, like many another holy mountain in which saints and dragons make their habitations in strange association with each other. In short, it was essentially a holy hill to which pilgrimages have been made annually on August 9, its sanctity going back probably to the time of the megalithic monuments at the base of the settlement.

Pagan Sanctuaries and Christian Shrines

Further to the west lies the city of Braga, the Roman *Bracara Augusta*, the capital of the *Callaici Bracarii*, where the cathedral is alleged to owe its origin to St. James of Compostela. About 4 miles to the south-east on a hill adjoining that on which stands the pilgrimage church of Bom Jesus do Monte, is the sanctuary of Sameiro in honour of the Virgin. In cutting a motor road to it a dolmen has been exposed in excellent preservation. So holy in fact was the hill in local estimation that the workmen engaged in erecting the Christian sanctuary on it had to be paid very liberally for undertaking such a perilous task.

These few examples from a mass of evidence which could be produced (restricted by considerations of time and space) may suffice to indicate that this region from the end of the Neolithic has been the scene of movements of peoples connected with a sacred tradition which has survived throughout the ages in the numerous pilgrimage centres and their associated cultus. As St. James was thought to have landed at Padròn and to

have found his final resting-place at Santiago, conveniently situated within easy access of the Atlantic and the ancient mining centres, so at the mouth of the Tagus in the cathedral at Lisbon the mortal remains of St. Vincent, the patron saint of sailors, are said to have been deposited in an urn, having been brought thither in a boat escorted by ravens, the descendants of which may still be seen in the courtyard ! This legend has a parallel in the story of St. Martha who, according to one version, was conveyed by ravens in a vessel without sails, oars, rudder or pilot, to a spot near Marseilles, ' where the Rhone pours into the Gallic sea.' From Palestine this mysterious craft made its way through the Adriatic and along the Mediterranean till it reached its final resting place where the immigrants, St. Martha, her sister Mary and their brother Lazarus, together with their companions, were well received in the province of Vienne. There they made their abode in a cave called St. Baune located in a huge rock. On the top of this rock stands the chapel of St. Pilon, and at the base a pile of stones near a desert inhabited by a dragon known as the Tarasque, thought by the Abbé Fouillon to have been a crocodile from Egypt. Having been subdued by St. Martha and then killed by the people of Tarascon (the home, incidentally, of the immortal Tartarin) which derives its name from the monster, its remains repose in the parish church. There, too, lie the relics of the saint in a small crypt called the Grotto (a Gallo-Roman tomb) under the western end of the north aisle. Originally they were buried, it would seem, in a rock-hewn cavity a little to the east of the place where they have been enshrined. For centuries the Grotto was one of the most famous sanctuaries in Provence. At it Clovis was alleged to have been cured of a mortal disease, and among notable visitors to the shrine were the Dauphin (afterwards Louis XI), Charles IX, Anne of Austria and Louis XIV.

If the St. Vincent legend could be connected with the tomb erected over the rude stone monument in the crypt of San Vincente at Avila, it would bring it into line with the St. Martha episode. The Spanish saint, however, appears to have been a contemporary of his Portuguese namesake, and after his martyrdom at Avila in 303, together with his sisters, SS. Sevina and Cristota, to have developed an independent cultus. On the other hand, the Acta of St. Leocardia places their death at Ebora, and if this is the equivalent of the modern Evora, the tradition thereby becomes connected with southern Portugal.

Be this as it may, the connexion between Christian shrines and their megalithic prototypes is so intimate in Iberian hagiology and folklore, often in association with a maritime legend concerning a mysterious voyage, usually from the eastern Mediterranean, that its significance and interpretation hardly can be in doubt. This is most apparent in the case of the cult of St. James at Compostela, where the tradition of sanctity appears to have lingered on from the Bronze Age until eventually Santiago became the centre of pilgrimage *par excellence* in medieval Europe, and attracted to the ancient megalithic site relays of votaries, especially from the Celtic fringes, following the same course pursued by their prehistoric and protohistoric predecessors along the ancient trade routes to Galicia.

Once sacredness is attached to a locality, object or person it does not readily lose its sanctity however much the original cultus may be transferred and transformed in course of diffusion from one supernatural being and his ritual to another—and a saint may be as popular as a god if his festival is kept in the same manner and more or less at the same season.

At the break-up of the Roman Empire Europe was immersed in a bewildering medley of mystery cults, confused mythologies, philosophic systems and popular traditions and customs derived from a great variety of sources, Latin, Celtic and Germanic. As the new unifying dynamic, Christianity recognised the value of this diverse and unco-ordinated material resulting from the intermingling of different racial, cultural, religious and historical traditions, and systematically set to work to use it for its own ends. Around its historic and symbolical figures—apostles, martyrs and confessors, as well as the Theotokos, the queen of saints—it developed a cultus, a legend and a sacred drama ;

often at first purely local in application and significance, but later given a wider reference as a result of cult contact. Pilgrimage normally was made to the place of burial of the martyrs and other heroes of the Faith, while the lives, exploits and adventures of the saints were systematically compiled by hagiographers after the cultus and its tradition had been established.

Where, as in the case of the earliest apostles, martyrs and evangelists, the actual place of interment was unknown, attempts were made to locate it at a spot that already had acquired a long reputation for sanctity. Thus, Compostela became the resting place of the mortal remains of St. James and Padròn the traditional scene of his evangelistic campaign ; while in a modified form with different associations the same tradition recurred south of the Minho, and persisted among the folk cultures of the Atlantic and Mediterranean littoral. In the process of adaptation to new and higher religious disciplines it underwent a good deal of re-interpretation, but so deeply laid was the maritime theme that it has remained a fundamental element throughout the spread and transvaluation of the ever-developing complex of belief and practice superimposed on a succession of cultures.[1]

References

Acta Sanctorum. VI, Julii. De S. Jacobo Majori, 6.

Allcroft, A. H. *The Circle and the Cross*. Vol. I (1927), Vol. II (1930).

Altamira. *Historia de España de la Civilizacion espanola*. Barcelona (1913), Vol. I.

Bégouen, M. *Compte Rendu de l'académie des Inscriptions*, XXIII (1912).

Borlase, W. C. *Age of the Saints* (1893).

Breuil, H, *Compte Rendu le l'académie des Inscriptions* (1940).

Childe, V. G. *The Dawn of European Civilization* (1939).

Coffey, G. *New Grange and Other Incised Tumuli in Ireland* (1912).

Ferreiro, G. A. *Historia de la Santo A. M. Iglesia de Santiago de Compostela*. Santiago (1898–1910).

Fergusson, J. *Rude Stone Monuments* (1872).

—— *History of Indian and Eastern Architecture*, Vol. I (1910).

Forde, C. Daryll. *American Anthropologist*, XXXII, No. 1, N.S. (1930).

Fleure, H. J. *Human Geography in Western Europe* (1910).

Gelmirez, Diego. *Historia Compostelana*, eleventh century (n.d.).

Greenwell, W. *British Barrows* (1877).

Howes, H. W. *Folk-Lore*, XXXVI (1925).

James, E. O. *Folk-Lore*, XLII (1932).

Johnston, P. M. *Sussex Archaeological Collections*, LV (1912).

Jones, P. Hartwell. *Celtic Britain and the Pilgrim Movement*, Y. Cymmrodor, Vol. XXIII.

Leeds, E. T. *Archaeologia*, LXX (1920).

Lopez, R. L. *Through Galicia to Santiago of Compostela*.

Macalister, R. A. S. *The Archaeology of Ireland* (1928).

Mortimer. J. R. *Forty Years' Researches in British and Saxon Burial Mounds in East Yorkshire* (1905).

Ovid, *Fasti*, ii.

Peake, H. *Folk-Lore*, XXX (1919).

Phillips, J. A. *A Treatise on Ore Deposits* (1896).

Powell, T. G. E. *Proceedings of the Prehistoric Society* (1938), N.S., Vol. IV, Pt. ii.

Verneau, R. *Les Grottes de Grimaldi*. Monaco, (1906) Vol. II.

Xanthoudides, S. *Vaulted Tombs of Mesara*. Liverpool (1924).

[1] Delivered Sept. 5, 1952.

THE NERVOUS CONTROL OF THE MOVEMENTS OF THE ALIMENTARY CANAL

Address by

PROF. R. C. GARRY

PRESIDENT OF SECTION I

> ' In the days when man's members did not all agree amongst themselves, as is now the case, but had each its own ideas and a voice of its own, the other parts thought it unfair that they should have the worry and the trouble and the labour of providing everything for the belly, while the belly remained quietly in their midst with nothing to do but to enjoy the good things which they bestowed upon it ; they therefore conspired together that the hands should carry no food to the mouth, nor the mouth accept anything that was given it, nor the teeth grind up what they received. While they sought in this angry spirit to starve the belly into submission, the members themselves and the whole body were reduced to the utmost weakness. Hence it had become clear that even the belly had no idle task to perform.' (LIVY, II, 32.)
>
> In 494 B.C. the plebeians rose in revolt and deserted Rome. Menenius Agrippa induced them to return by recounting the above fable.

WHEN an organism is highly developed the various systems within the organism must be subject to control. Such control can be exerted both by humoral agencies and by the nervous system. It is possible to envisage the form this control ought to take if optimal functional integration is to be achieved within the organism.

(i) The system should be integrated within itself, the various parts should act in harmony with one another.

(ii) But a system, however well co-ordinated, is not an end in itself. It is part of a larger whole, the organism. We therefore will expect to find influences at work bringing the system into subjection to the economy of the entire body.

(iii) We will expect to find, too, some device to enable the system to react on its parent organism.

Brocklehurst (1950) in his Presidential Address to Section I in 1950 dealt with the rôle of humoral agencies in the control of the alimentary system. This, then, is to be an attempt to give a companion description of the part played by the nervous system in control of the movements of the gastro-intestinal tract.

Over fifty years ago Bayliss and Starling (1899) wrote, and it is the first sentence of their classical paper ' The movements and innervation of the small intestine,' ' On no subject in physiology do we meet with so many discrepancies of fact and opinion as in that of the physiology of the intestinal movements.' And these two did solve, to their own satisfaction and to that of their contemporaries, the outstanding problems of the innervation of the gut. Their findings, little altered, are the basis of most text-book descriptions at the present day. Yet close reading of their papers, and of many articles published subsequently by others, reveals no little doubt and uncertainty. The time is ripe for a review of the present position.

1. THE INTRINSIC INTEGRATION OF THE MOVEMENTS OF THE ALIMENTARY CANAL

Magnus (1904) found that the small intestine, *in vitro*, showed rhythmical activity and tone. He and later workers claimed that such movements occur in portions of gut devoid of nerve cells. It is

very doubtful, in the light of modern histological observations, if one can by simple dissection remove nerve cells, let alone all nervous elements, from the intestine. Yet the gut continues to show rhythmical activity in the presence of nerve poisons and also after intervals of time which one would expect to be lethal to nervous tissue.

What are these movements exhibited by the Magnus preparation of the gut *in vitro*? What do they correspond to in the intact gut in the intact animal?

Bayliss and Starling (1899) saw three types of movement in the small intestine when the abdomen of an anaesthetised dog had been opened under a saline solution. There were *pendulum movements*, a swaying to and fro of loops of the gut floating in the solution. There were slight *rings of constriction* which usually moved somewhat rapidly caudad, although they did, on occasion, move craniad. Lastly, there were *peristaltic movements*, powerful zones of constriction preceded by inhibition, which invariably moved caudad. It is to be doubted, however, if Bayliss and Starling regarded pendulum movements as a separate entity. In a later paper (1900–01), they describe 'rhythmic contractions giving rise to the pendulum movements and running rapidly down the intestine.' In the summary of this paper they refer to 'Myogenic rhythmic contractions ("pendulum movements").'

In 1917 Trendelenburg was able to show peristalsis *in vitro* in the gut of guinea-pigs. When an excised portion of the small intestine was distended it started to shorten until, at a definite critical pressure, a constriction appeared at the cranial end of the segment. A wave of contraction then swept rapidly caudad expelling the contents of the gut. This peristaltic wave was accompanied by lengthening of the intestine. Such peristaltic movements *in vitro* show rapid fatigue, disappear long before simple rhythmical movements cease, and are abolished by nicotine which does not impair the 'pendulum' movements.

We now know that these peristaltic waves in the Trendelenburg apparatus are abolished by cocaine, by *d*-tubocurarine and by hexamethonium (Burn, 1952).

The remarkable self-sufficiency of the gut is also shown by the 'visceral organism' of Markowitz and Essex (1930).

Cannon (1911–12) described 'tone rings' in the gut. These are local stationary circumscribed constrictions which give off wave after wave of contraction which may move caudad (katastalsis) or craniad (anastalsis). The incisura angularis, the prepyloric sphincter, is such a 'tone ring' giving rise to the pyloric movements in the stomach. In the large bowel a similar 'tone ring' may be established and, in this case, the waves of contraction move craniad towards the caecum. This is the 'antiperistalsis' of the colon. Similar 'pulsating tonic rings' appear in the small intestine and they are not abolished by nicotine. It is attractive to imagine that the waves of contraction seen by Bayliss and Starling are the waves which Cannon believed to arise from 'tonic rings.'

In 1902 Cannon detected still another movement in the small intestine of unanaesthetised cats. He added bismuth subnitrate to their food so that the contents of the gut would be opaque to X-rays. He then watched the intestinal movements on a fluorescent screen. At intervals along the small intestine constrictions simultaneously developed, remained for some time and then disappeared while a new series of constrictions made its appearance halfway between the previous constrictions. This co-ordinated sequence of activity is *segmentation*.

The gut then seems to show four types of activity :—

(i) Rhythmical movements *in vitro* which may, or may not, represent an activity in the intact animal. Bayliss and Starling regarded them simply as a response of the smooth muscle to stretch. For this activity 'pendulum movements' would be a very good name but, unfortunately, it has been applied to a hypothetical type of gut movement *in vivo*.

(ii) Waves of contraction which usually move caudad and which may arise in pulsating tonic rings.

(iii) Segmentation with its alternating constriction and dilatation of the small intestine.

(iv) Peristalsis, a wave of contraction moving caudad and preceded by inhibition.

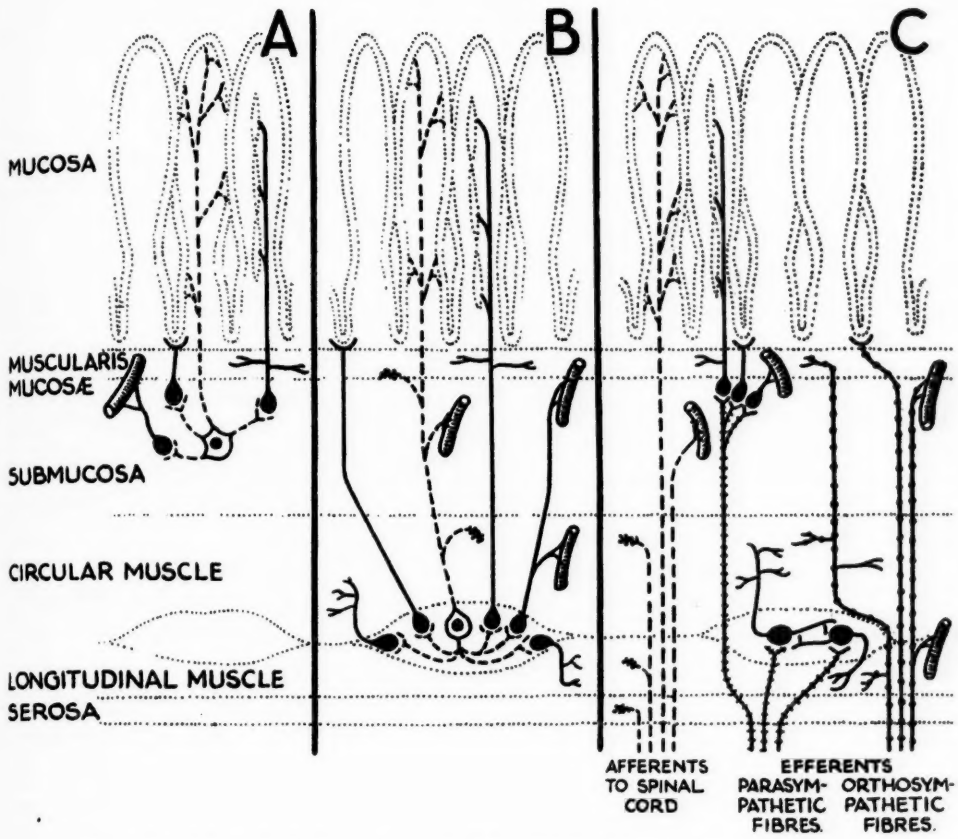


Figure 1. A hypothetical arrangement of the nervous elements within the gut wall.

A. *Local mucosal reflexes.* Ingoing pathway from the epithelium to Meissner's plexus : outgoing pathways to secretory cells of the crypts, to the blood-vessels, to the muscularis mucosae and to the smooth muscle in the villi.

B. *Intrinsic reflex, local and spreading, involving entire bowel wall.* Ingoing pathway from the epithelium to Auerbach's plexus : outgoing pathways to the muscle coats, to blood-vessels and to mucosal structures as in A.

C. *Spinal reflexes.* Ingoing pathway from the epithelium, blood-vessels, muscle coats and serosa to the spinal cord : outgoing pathways by ortho- and para-sympathetic outflows to the gut.

H. S. D. GARVEN.

The nature, control and integration of these movements, however, cannot be further discussed without some description of the nervous structures known to be present in the gut, Fig. 1.

In 1927 Catherine Hill published a comprehensive paper on the histology of the enteric plexuses. She somewhat too skilfully crystallised our views on the matter, for no material alteration or modification of her presentation has found its way into current physiological thought.

By application of nicotine early workers soon showed that the orthosympathetic outflow is interrupted in collateral ganglia remote from the gut and that non-myelinated post-ganglionic axons run thereafter directly to a small proportion of the muscle cells belonging to both muscular coats of the intestine. The axons of the parasympathetic outflow, on the other hand, run directly to the gut without interruption. There they synapse with the cells of Auerbach's myenteric plexus, Fig. 1c. Langley (1922) postulated the existence of two types of cell in the myenteric plexus. One type receives parasympathetic axons with either an

inhibitory or motor action. Such cells are few in number. The other type serves to distribute either the motor or inhibitory impulses of parasympathetic origin. These are ' local cells,' and are numerous. Ambache (1951), as a result of very recent work, goes a step further. He also finds evidence that the ' local cells ' of Langley are of two types, in modern terminology cholinergic and adrenergic. Moreover, he ventures the suggestion that the cholinergic local cells send their axons craniad, the adrenergic cells caudad.

The position is still further complicated by the claims of Boeke (1940). He does not believe that a tiny end-knob, which is present on only one in about 100 smooth muscle cells, can liberate a sufficiency of the transmitter which we now believe to be set free at neuro-effector junctions. He therefore postulates the existence in the gut, as elsewhere, of an ' interstitial cell ' net which acts as the real termination of the autonomic system, „ eine Umvertungstelle der Erregung."

However this may be, there do seem to be grounds for believing that the myenteric plexuses form a rather elaborate nerve net which may be polarised. What part does this nerve net play in control of intestinal movements ?

The ' pendulum movements ' *in vitro* and the waves of contraction which usually move caudad, types (i) and (ii) above, may be truly myogenic. Co-ordination of these activities can then be only by direct spread of excitation or of inhibition through the quasi syncytium of the gut musculature.

Cannon (1912) regarded segmentation, type (iii), as a local response to distension, the response depending on the integrity of nervous elements locally in the gut. He was careful to point out that a reflex, in the true sense of the word, is not involved.

The majority of physiologists are now agreed that true peristalsis, the type (iv) movement, involves contraction craniad and inhibition caudad to the site of stimulation. Bayliss and Starling accepted such co-ordination as evidence for the existence of a nervous reflex in the myenteric plexus. Cannon (1912) says ' Probably in the usual (reflex) peristalsis of the small intestine the same neuromusculature is involved as in segmentation, but subject to superintending and co-ordinating neurons.' Cannon gave the name ' diastalsis ' to this complex movement. Gasser (1926), referring to peristalsis elicited *in vitro* by the Trendelenburg technique, says ' It is difficult to conceive a co-ordinating mechanism for this interplay not involving an intercalated nervous element.'

However, several authorities will have none of this. To Johnson (1924–25) the myenteric plexus is merely a diffuse collateral ganglion. Carey (1921) studied the cyto-architecture of the coats of the gut. He came to the conclusion that the disposition of the muscle coats is such that distension of the gut, which leads to contraction locally, must bring about constriction craniad and dilatation caudad. He believes it is quite gratuitous to postulate the existence of a nervous reflex. Bayliss and Starling suggested that the preponderance of descending waves of contraction, type (ii) activity in the present classification, ' may be due to the higher excitability of the fibres at the duodenal end of the gut.' Alvarez (1949), as is well known, used this conception of a gradient of activity from the cranial to the caudal ends of the gut to explain the onward passage of the gut contents. He also does not think it necessary to invoke the help of a reflex in the myenteric plexus.

There it is probably wise to leave the controversy. As Occam said, one should not postulate the existence of factors without the necessity to do so. But surely it is equally wrong to ignore factors undoubtedly present. And it is difficult to believe that the myenteric plexus, as we now know it, is not a factor in reflex control of at least certain of the intrinsic activities of the gut. But one must admit that a vital piece of evidence is still missing. A reflex postulates the existence of a receptor, an ingoing path, a ' centre,' an outgoing path and an effector. There are no highly organised receptors in the gut, and ingoing axons, according to Catherine Hill, pass straight through from the mucosa or blood-vessels without synapse in the myenteric plexus, Fig. 1c. In spite of the absence of specialised receptors we now know that stimuli, both

normal and abnormal, when applied to the mucosa, submucosa and muscle coats of the gut, send impulses into the central nervous system. If the existence of collaterals from the ingoing axons to the myenteric plexus could only be shown then there would be the neural basis necessary for local reflexes within the intestine itself. Alternatively, stimulation of the gut may, in some other fashion, elicit nerve impulses within the myenteric plexuses, possibly through the medium of local 'sensory' neurones such as are depicted in Fig. 1A and B.

There is another activity which may have a bearing on this general problem. Although the intestinal villi respond by increased activity to some humoral agent produced during digestion, a single villus can also contract, and smartly too, when the mucous membrane close to the base of the villus is touched. This response is said to depend on the integrity of nervous elements in the submucosa, Fig 1A.

The intestinal movements so far considered are regional and controlled locally. Is there evidence for a more general control, a control still intrinsic and divorced from the extrinsic innervation?

Macewen in 1904 noted that taking food into the human stomach promptly led to increased activity of the caecum, colon and possibly of the ileum. The response 'followed the ingestion of food too quickly to permit of the peristalsis acting through the whole length of the small intestine. It was, therefore, regarded as probably due to reflex action.' For many years this phenomenon has been known as the gastro-colic reflex.

Douglas and Mann (1939, 1940) and Douglas (1941) re-examined this phenomenon in dogs. Eating, placing food directly into the stomach or into the duodenum through a fistula, caused, within a few seconds, a marked and prolonged increase in the activity of the ileum. This rapid response persisted after section of both vagus nerves, but could not be obtained from Thiry-Vella loops with intact extrinsic innervation. On the other hand, section of the gut, with immediate resuture, did not abolish the effect. So Douglas and Mann suggest that the gastro-colic reflex is no more than a rapidly propagated local response of the gut to distension. It is rather like the jogging wave which passes down a line of railway trucks during shunting. But Douglas and Mann report one rather odd feature. While the distal colon showed a gastro-colic response, the proximal colon did not do so. There is a gap, then, in the line of trucks and we do not know what bridges the gap. Nevertheless, there is little doubt that 'a wave of activity passes down the whole length of the bowel when food is first taken after a fast.' For this phenomenon Douglas (1948) suggests 'post-prandial excitation of the intestine.'

Sir Arthur Keith in 1915 put forward the interesting suggestion that the gut may possess intestinal 'pacemakers' on analogy with the sinu-atrial node of the heart. In the myenteric plexus he described a 'third element,' a 'neuro-muscular tissue' which formed 'nodes' at the gastro-oesophageal junction, at the cranial portion of the duodenum, at the ileo-colic junction and also at certain sites in the large intestine. Douglas too (1949) believes that the duodenum acts as a rhythmogenic centre. If a portion of the jejunum is transplanted to the ileum the frequency of the jejunal contractions becomes less. The duodenum thus seems to drive the jejunum to show a frequency of contraction greater than that natural to the jejunum. This conception, like that of Keith, postulates the existence of an intrinsic co-ordinating system, be it the visceral muscle itself, a specialised conducting system, or the nerve net of the myenteric plexus.

It is only too obvious that our knowledge of the nature and control of the intrinsic powers of the intestinal tract is defective. We can only say with certainty that the gut is a thoroughly self-sufficient structure exhibiting no small degree of co-ordination within itself. The relative importance of muscular and nervous elements in this co-ordination is still undecided.

2. Control of the Gut within the Body

The gut receives its nerve supply from the Autonomic Nervous System, Fig. 2. This at least partially autonomous system is divided into (*a*) Orthosympathetic and (*b*) Parasympathetic portions. Most of the

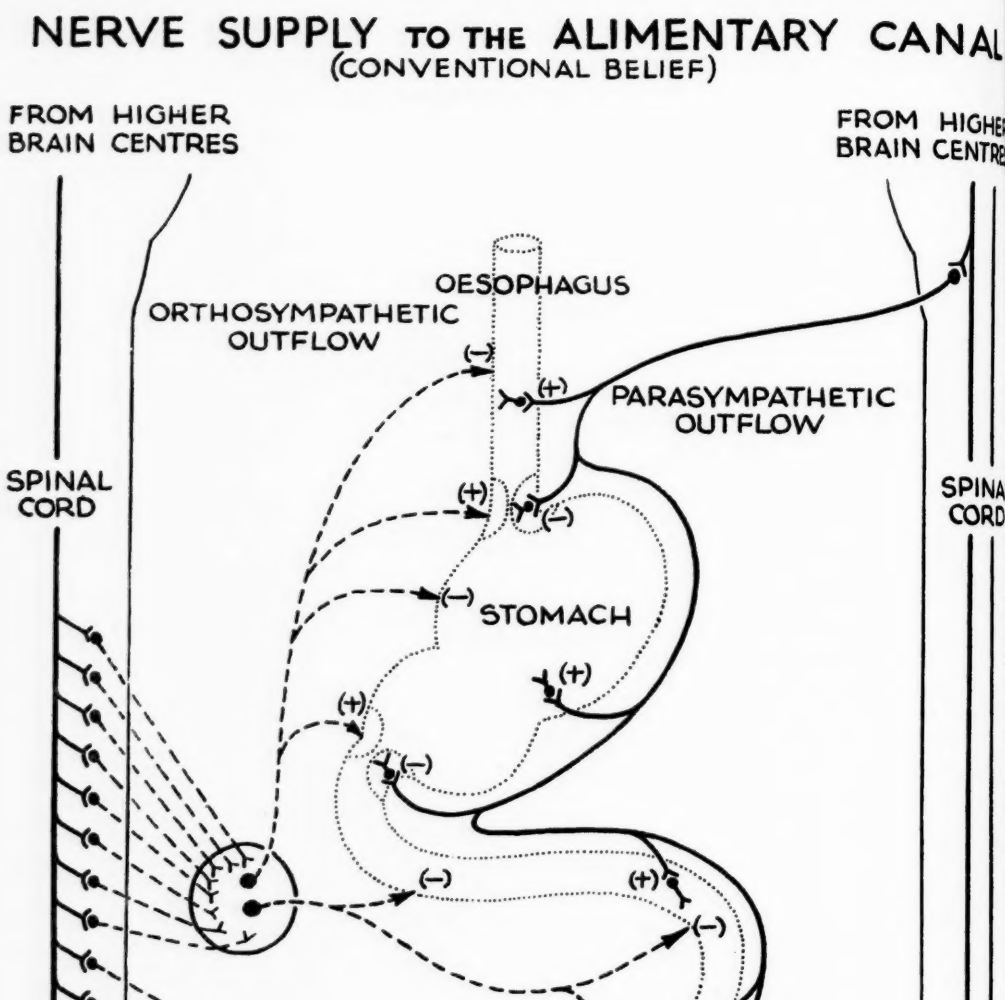


Figure 2. A simple diagram of the nervous outflow to the alimentary canal.

gut is thought to receive its parasympathetic supply through the vagus nerves, although the caudal region is supplied by the pelvic nerves which also belong to the parasympathetic outflow.

It is most attractive, and dangerously satisfying mentally, to see antagonistic controls in these two outflows to the gut. Probably the real state of affairs is much more complicated and not nearly so clean cut. It is certainly an ominous fact that there has been little change in our fundamental views since the days of Gaskell (1916) and of Langley (1921). We are no nearer understanding of the fact that stimulation of these outflows may cause contraction of smooth muscle at one site while, close by, smooth muscle, to all appearances the same, is inhibited.

Essentially only three techniques are available to the physiologist who wishes to unravel the functions of the nervous outflows to the gut. He may cut the outgoing nerves and then stimulate artificially the peripheral ends of the cut nerves. Alternatively he may simply sever the outflows and observe the effects following on absence of normal outgoing impulses. Both these methods have fundamental defects. Anatomical nervous outflows may not be ' pure,' they may contain axons belonging to another outflow and the inevitable operative procedure may create an unphysiological state. Less objectionable, perhaps, is to observe the effect of administration of the substances thought to be the humoral transmitters of the autonomic system. Substances known to act upon these transmitters may also be administered and the effect noted.

Control by the Orthosympathetic Outflow.—It is generally believed that the orthosympathetic outflow is inhibitor to the gut and motor to the sphincters. In the words of Elliott (1905) ' . . . if quiet lodgement of the contents be facilitated by the presence of sympathetic inhibitor nerves to the body of the viscus, there will also be sympathetic motor nerves to the sphincter closing the exit.'

Probably the most satisfactory evidence comes from the surgeons who obtain the expected result in patients when they interrupt the orthosympathetic outflow on the hypothesis that it is inhibitor to the gut and motor to the sphincters.

Yet certain niggling doubts remain. Artificial stimulation of what are thought to be orthosympathetic outflows to the stomach may produce motor effects (McCrea and McSwiney, 1928), and adrenaline in the guinea-pig can cause contraction of the small intestine close to the ileo-colic sphincter (Munro, 1951). Moreover, it is said that the frequency with which the stimulation is applied may affect the nature of the response (Finkleman, 1930).

Nevertheless, we may venture to ascribe three functions to the orthosympathetic outflow. (1) It may ' allow filling and time for absorption ' (Learmonth and Markowitz, 1930). (2) It may inhibit gastro-intestinal activity when blood is diverted from the viscura, say during muscular exercise. (3) When mischief is afoot in the abdominal cavity the orthosympathetic outflow inhibits the gut. This inactivity may prevent exacerbation of the pathological process.

Control by the Parasympathetic Outflow.—On the theory of autonomic balance one would expect the parasympathetic outflow to be motor to the gut and inhibitor to the sphincters. Experimental results are far from being in keeping with this view. No one has been able to find a definite inhibitor outflow to the ileo-colic sphincter. McCrea, McSwiney and Stopford (1925) go as far as to compare the action of the vagus nerve on the stomach to the action of the same nerve on the heart. Stimulation of the peripheral end of the cut vagus nerve initially causes inhibition which is followed by ' vagus escape ' and an augmentation of the stomach movements. McCrea and McSwiney (1928) find ' no evidence of any direct antagonism between the vagus and splanchnic nerves.' McCrea and Macdonald (1928–29) state that the ' theory of vagus nerve being excitor and splanchnic inhibitor is untenable.'

These workers all emphasised the importance of what they called the ' peripheral mechanism.' When the stomach is active, stimulation of the orthosympathetic and also of the parasympathetic outflow will tend to cause inhibition. When the stomach is inactive, whatever outflow is stimulated, the result will tend to be motor.

There is no reason to doubt the accuracy

of these findings. Probably our understanding of the nature of the extrinsic innervation of the gut is at fault, we are not asking the questions we think we are asking when we cut nerves which supply the gut and then stimulate the peripheral ends artificially. More in keeping with the conventional hypothesis is the observation that, when the vagus nerves are cut, there is delay in emptying of the stomach. When the pelvic nerves are cut, the movements of the large bowel are impaired and reflex defaecation is abolished (Garry, 1934).

Gregory (1947) showed that activity in the jejunum, induced by administration of apomorphine, was abolished by cutting the vagus nerves. Later, in 1950, he showed that the act of feeding, and also the presence of food in the stomach, increased the activity of the jejunum. Moreover, he found, rather unexpectedly, that this increased activity delayed the onward movement of fluid flowing through the jejunum. Cutting the nerves to the gut abolished this effect. It is difficult to reconcile such findings with the work of Douglas and Mann quoted above.

In view of these and other observations it is very difficult indeed to ascribe with confidence a definite function to the parasympathetic outflow to the gut. We do know, however, that the integrity of the pelvic nerves is necessary for defaecation. On the hypothesis that the parasympathetic is motor to the gut and inhibitor to the sphincters the idea has grown that the parasympathetic promotes translation of the contents of the gut caudad. But Gregory believes that increased activity delays translation to allow time for absorption. As noted above this very function has been attributed to the orthosympathetic outflow since it inhibits gut activity. The answer is to be found in further work.

Control at the Cranial and Caudal Extremities of the Gut.—The cranial end of the oesophagus has walls composed of striped muscle, the extent of this musculature depending on the species of animal. Division of the vagus nerves causes permanent paralysis of the striped muscle (Cannon, 1907). The caudal end of the oesophagus has smooth muscle. In this case, section of the vagus nerves causes only transient paralysis of the smooth muscle, and, for a time, the cardiac sphincter does not relax normally to allow food in the oesophagus to enter the stomach. When food is swallowed, the walls of the stomach most accommodatingly show ' receptive relaxation.' This response depends on the integrity of the vagus nerves (Cannon and Lieb, 1911).

In suckling ruminant animals there is a functional extension of the oesophagus. In order to bypass the undeveloped forestomachs an oesophageal groove is formed by contraction of the walls of the stomach. This groove conveys the milk directly into the abomasum, the real stomach. The groove appears only during the act of suckling and its formation depends on an intact nerve supply.

Vomiting is a complex reflex act involving the striped muscle of the abdominal wall and of the diaphragm as well as the smooth muscle of the gut. The vagus nerves are thought to be the chief outgoing pathways to the gut in this reflex although a patient with severed vagus nerves can still vomit (Illingworth, 1952).

At the caudal end of the gut the pelvic nerves are motor to the colon and rectum and inhibitor to the smooth internal anal sphincter. The orthosympathetic outflow has the opposite action. This generalisation has been challenged by Lawson and Templeton (1932).

The caudal end of the gut seems to have lost much of its autonomy. Division of the pelvic nerves, or destruction of the portion of the spinal cord from which these nerves arise, leads to long-lasting paralysis of the caudal end of the gut. Yet, even in man, the caudal end of the gut may regain much of its intrinsic power after destruction of the spinal cord (Head and Riddoch, 1917 ; Denny-Brown and Robertson, 1935).

It would seem that the function of the parasympathetic outflow at the termination of the gut is to evacuate the contents of the large bowel, while that of the orthosympathetic outflow is to inhibit colonic activity, possibly when the organism as a whole faces an emergency.

Since we know now that there is a centre for the autonomic nervous system in the hypothalmus, in theory at least the

entire length of the intestine may be subject to impulses arising in the brain. Undoubtedly, too, the cortex of the brain may influence the hypothalamus. In this way may be explained the sporadic references in the literature to effects on the gut when the highest centres of the brain are stimulated. Such observations, however, are far from justifying the view that the highest nervous centres are constantly about the business of controlling the gut or the viscera in general.

3. The Influence of the Gut on the Body—Ingoing Nervous Impulses

One is faced straight away with the fact that no elaborate nervous receptors have been described in the substance of the alimentary canal. In man the intestine can be cut, burned and crushed without giving rise to sensation. Nevertheless, as everyone knows, sensation can arise from the gut. Colic, which need not arise in the colon, is *the* example of such sensation. The position is still further confused by controversy over the nature of sensation from the viscera. Is there a true visceral sensation or is the sensation ' referred ' to the surface of the body ? We seem to be now veering towards the belief that there is a true visceral sensation. However this may be, there is no doubt that nerve impulses do travel from the gut to the central nervous system. This was demonstrated by McDowall (1925) who noted that distension of the intestine caused dilatation of the pupil of cats anaesthetised with chloralose. With this technique McSwiney and his pupils traced the complicated nervous pathways into the central nervous system. Probably nerve impulses are constantly passing into the central nervous system from the active gut but rarely reach consciousness. When the contracting gut meets a resistance in its lumen which it cannot overcome then the impulses reach consciousness to elicit the pain of colic.

Ingoing nervous impulses from the intestine may play a part in reflex co-ordination of the activity of the alimentary canal. Often impulses from the gut and from other viscera bring about inhibition of the gut reflexly (Youmans, 1949). It is possible that such reflexes are protective.

At the extremities of the intestinal tract we have rather better information concerning the nature and importance of ingoing nerve impulses.

Nausea and vomiting can be evoked by stimuli applied to the pharynx, the oesophagus, the stomach, indeed to many parts of the alimentary canal. We have no histological evidence for specific nerve endings initiating the impulses for this complicated protective act. After division of the vagus nerves a human being can still vomit (Illingworth, 1952), so some other pathway must be followed by the ingoing nerve impulses.

Hunger, the real gnawing pangs of hunger, are due to powerful contractions of the empty stomach (Carlson, 1916). The sensation of hunger ceases when the stomach is filled and settles down to the less stormy activities of gastric digestion. Hunger pangs can be relieved even when the contents of the stomach have no nutritive value. The nerve impulses which give rise to a sensation of hunger probably travel inwards to the nervous system along the nerves used by the orthosympathetic outflow. For once, however, there can be little doubt about the function of a nervous inflow.

From observations on human beings with an opening between the stomach and the abdominal wall we know that vague sensations, crudely localised, can be elicited from the stomach (Wolf and Wolff, 1943). Certain of these sensations persist after division of the vagus nerves, so some other ingoing pathway must be available (Illingworth, 1952). The pain of peptic ulceration, the discomfort of indigestion and of heartburn must also be elicited by ingoing nerve impulses. We know little about their causation and the route they follow but we can attribute a protective function to these nerve impulses.

At the caudal end of the gut there is some certainty about the routes followed by the ingoing nerve impulses. Defaecation, like vomiting, is an elaborate reflex involving the large intestine and skeletal musculature. From the terminal portion of the large intestine, the rectum, the presence of faecal matter evokes initially a characteristic sensation and initiates the reflex evacuation of the colon. The nerve impulses travel inwards along the pelvic

nerves, the nerves used by the parasympathetic outflow (Garry, 1934 ; Goligher and Hughes, 1951). Once more information about the nature of receptors for this special inflow of nerve impulses is lacking. Goligher (1951) believes that the receptors are situated in the mucosa of the rectum since removal of the mucosa, even with retention of the muscular wall of the rectum, abolishes the sensation elicited from the rectum.

From the large intestine cranial to the rectum distension does not elicit defaecation, but pain may be evoked and the impulses pass inwards along pathways used by the orthosympathetic outflow.

Our knowledge concerning the routes followed by, and the nature and functions of, the impulses which travel inward from the gut is grossly defective. There are, however, signs of growing interest in this field.

Conclusion

The object of this review is to show the alimentary canal as an integrated system controlled by, and influencing in its turn, the organism as a whole. How far performance has lagged behind intention is only too obvious. There are woeful gaps in our knowledge. Observed facts, which there is no reason to doubt, will not fit into hypotheses inherited from the past. Our only recourse is to persevere, to accumulate still more facts, in the hope that ultimately we shall see the picture whole.[1]

Fig. 1 is reproduced by courtesy of Dr. H. S. D. Garven.

References

Alvarez, W. C. (1949). *An introduction to gastroenterology.* 4th ed. London : Heinemann.

Ambache, N. (1951). *Brit. J. Pharmacol.*, **6**, 51.

Bayliss, W. M. & Starling, E. H. (1899). *J. Physiol.*, **24**, 99.

Bayliss, W. M. & Starling, E. H. (1900–01). *J. Physiol.*, **26**, 125.

Boeke, J. (1940). *Problems of Nervous Anatomy.* London : Humphrey Milford.

Brocklehurst, R. J. (1950). *Advanc. Sci., Lond.*, **7**, 212.

Burn, J. H. (1952). *Practical Pharmacology.* Oxford: Blackwell.

Cannon, W. B. (1902). *Amer. J. Physiol.*, **6**, 251.

Cannon, W. B. (1907). *Amer. J. Physiol.*, **19**, 436.

Cannon, W. B. (1911–12). *Amer. J. Physiol.*, **29**, 238.

Cannon, W. B. (1912). *Amer. J. Physiol.*, **30**, 114.

Cannon, W. B. & Lieb, C. W. (1911). *Amer. J. Physiol.*, **29**, 267.

Carey, E. J. (1921). *Anat. Rec.*, **21**, 189.

Carlson, A. J. (1916). *The Control of Hunger in Health and Disease.* Chicago : Univ. of Chicago Press.

Denny-Brown, D. & Robertson, E. G. (1935). *Brain*, **58**, 256.

Douglas, D. M. (1948). *J. Physiol.*, **107**, 472

Douglas, D. M. (1949). *J. Physiol.*, **110**, 66.

Douglas, D. M. & Mann, F. C. (1939). *Amer. J. dig. Dis.*, **6**, 434.

Douglas, D. M. & Mann, F. C. (1940). *Amer. J. dig. Dis.*, **7**, 53.

Elliott, T. R. (1905). *J. Physiol.*, **32**, 401.

Finkleman, B. (1930). *J. Physiol.*, **70**, 145.

Garry, R. C. (1934). *Physiol. Rev.*, **14**, 103.

Gaskell, W. H. (1916). *The involuntary nervous system.* London : Longmans, Green & Co.

Gasser, H. S. (1926). *J. Pharmacol.*, **27**, 395.

Goligher, J. C. (1951). *Ann. Roy. Coll. Surg. Eng.*, **8**, 421.

Goligher, J. C. & Hughes, E. S. R. (1951). *Lancet*, I., 543.

Gregory, R. A. (1947). *J. Physiol.*, **106**, 95.

Gregory, R. A. (1950). *J. Physiol.*, **111**, 119.

Head, H. & Riddoch, G. (1917). *Brain*, **40**, 188.

Hill, Catherine J. (1927). *Philos. Trans. B.*, **215**, 355.

Illingworth, C. F. W. (1952). *Personal communication.*

Johnson, S. E. (1924–25). *J. comp. Neurol.*, **38**, 299.

Keith, A. (1915). *Lancet*, II, 371.

Langley, J. N. (1921). *The autonomic nervous system.* Cambridge : Heffer & Sons.

Langley, J. N. (1922). *J. Phyisol.*, **56**, 39 P.

Lawson, H. & Templeton, R. D. (1932). *Amer. J. Physiol.*, **100**, 362.

Learmonth, J. R. & Markowitz, J. (1930). *Amer. J. Physiol.*, **94**, 501.

Macewen, W. (1904). *Brit. Med. J.*, II, 873.

McCrea, E. D'A. & Macdonald, A. D. (1928–1929). *Quart. J. exp. Physiol.*, **19**, 161.

McCrea, E. D'A. & McSwiney, B. A. (1928). *Quart. J. exp. Physiol.*, **18**, 301.

McCrea, E. D'A., McSwiney, B. A. & Stopford, J. S. B. (1925). *Quart. J. exp. Physiol.*, **15**, 201.

McDowall, R. J. S. (1925). *Quart. J. exp. Physiol.*, **15**, 177.

Magnus, R. (1904). *Pflüg. Arch. ges. Physiol.*, **102**, 123.

Markowitz, J. & Essex, H. E. (1930). *Amer. J. Physiol.*, **92**, 205.

Munro, A. F. (1951). *J. Physiol.*, **112**, 84.

Trendelenburg, P. (1917). *Arch. exp. Path. Pharmak.* **81**, 55.

Wolf, S. & Wolff, H. G. (1943). *Human gastric function : an experimental study of a man and his stomach.* London : Oxford University Press.

Youmans, W. B. (1949). *Nervous and neurohumoral regulation of intestinal motility.* London : Interscience Publishers, Ltd.

[1] Delivered Sept. 8, 1952.

THE ASSESSMENT OF PERSONALITY

Address by

PROF. P. E. VERNON

PRESIDENT OF SECTION J

PRESIDENTIAL Addresses, to this Section at least, usually fall into one of two types. The first is the broad, urbane, survey of the historical development of the subject and of its rôle in our culture. The other type is the more detailed discussion of some special field in which the speaker is particularly interested. The first is often witty, the second technical. My Address clearly belongs to the second class. But though I cannot aspire to wit or urbanity, I will at least try not to be unduly technical.

By personality I mean, simply, what kind of a person is so-and-so, what is he like? This is not quite the same as the colloquial usage, whereby personality refers merely to the extent to which a person impresses or attracts other people. Rather it connotes the whole of a person's outstanding characteristics, his emotional and social traits, interests and attitudes, which distinguish him from other personalities. You can call this temperament, or character, or disposition, if you prefer. I shall not attempt a more precise definition, but will discuss the theoretical nature of personality later.

Now though I have been working in this field of personality, on and off, for twenty-five years, I cannot honestly claim that I, or other psychologists, have made much progress in the practical task of assessing personality simply and accurately. We fully realise the importance of taking personality factors into account in guiding or selecting children and adults into suitable educational or vocational careers. But we are certainly not in a position to provide the teacher, the Youth Employment Officer or the personnel official, with a straightforward battery of personality tests at all comparable to our tests of aptitudes and attainments. At the same time there have been considerable advances, especially in experimental work into problems of personality. We do possess the tools for investigating, for example, the origins of anti-semitic and other prejudices, the influence on children's personalities of education in progressive as contrasted with formal schools, the effects on their development of different kinds of home atmosphere or of treatment during infancy, the personality changes brought about in adults by such brain operations as leucotomy, or by electro-shock- or psycho-therapy, and so on. I am not claiming, of course, that we have complete answers to such questions; but the results obtained by scientific research are certainly worth more than those reached by the kind of journalism which still bulks so largely in most books on medical, social, educational and industrial psychology. Unfortunately such research is exceedingly elaborate, and it can be undertaken only by skilled psychologists, preferably working in a team with psychiatrists, sociologists and statisticians. Thus although it does much to extend our knowledge of personality psychology, and to indicate the inadequacies of the tests or other methods that have commonly been adopted in personality assessment, it gives little positive help in the provision of more accurate methods for everyday use.

THE INTERVIEW

What then are the methods which the layman or the psychologist normally uses for diagnosing the personality qualities of an individual whom he is trying to guide or to select? How far are they defective, and why? The main methods are, first, casual observation of the individual's behaviour at school or in employment or elsewhere, and, secondly, the interview. Judgments of personality in

an interview are based partly on observations of manner, expression and speech, partly on a more or less systematic inquiry into the individual's past history and present aims. Descriptions or testimonials written by acquaintances may also be consulted. Now there has been a great deal of experimental investigation of the selection interview, most of which shows it to be extremely weak both in reliability and validity. By low reliability I mean that two interviewers seeing the same candidates independently tend to arrive at very different conclusions about them. The correlation coefficient, which gives a statistical index of their agreement, is typically around 0·5, though occasionally much better figures up to 0·8 are reached when trained interviewers analyse carefully beforehand just what qualities they are looking for. The more typical coefficient of 0·5 may be illustrated as follows. Suppose two grammar-school headmasters each interviewed 100 11-year candidates and tried to pick the best 20 for grammar-school education ; they would agree only on 9 of their choices ; each would accept 11 whom the other would reject.

Hartog and Rhodes's investigation, described in *An Examination of Examinations* in 1935, is worth recalling, since it studied a representative procedure for selecting higher civil servants. Sixteen candidates were interviewed in the normal way by two distinct boards of experienced interviewers. The boards had access to examination marks and past records, and each awarded a mark for ' value of the candidate's personality for the Home Civil Service.' Members within each board showed fair agreement, but the correlation between the two boards was only 0·41, and discrepancies in marks ranged up to 31 per cent. It is not only civil servants, employers and teachers whose judgments are inconsistent. Some studies have shown that when two or more psychiatrists interview cases of nervous breakdown or mental illness, they disagree in their psychiatric diagnoses of about half the cases (the precise figure depends on the heterogeneity of the patients and the coarseness of the classification).

Consider next the predictive value of judgments made in interviews. During the war, recruits were given on entry a short series of group tests of intelligence, education, mechanical and other abilities ; and they filled in a questionnaire describing their previous education, employment history, relevant experience and interests. Each recruit was then interviewed for about 15 minutes by a Personnel Selection Officer (P.S.O.), and a recommendation was made for a suitable employment, in the light of all the evidence. These P.S.O.s, it should be noted, were not psychologists, but serving officers who had been carefully chosen and trained by psychologists. In several experiments, the success or failure of recruits recommended for different jobs was followed up, and it was commonly found that predictions by the average P.S.O. were less accurate than predictions based on the group tests alone. But there were large differences. For example, among a hundred or more electrical mechanics selected for the Royal Navy by one P.S.O., 37 per cent. failed to qualify at the job ; whereas among an equally big group selected by another only 18 per cent. failed. Clearly many P.S.O.s, although they took account of test scores, previous history and interests, and though they tried to assess the recruits' personality qualities, did worse than if they had never seen the recruits at all, but had mechanically recommended those with the best test scores.

A similar research on selection of students at a college of London University has been reported by Dr. Himmelweit. The students' marks on a short entrance examination, and scores on a number of psychological tests, together with assessments of intellectual promise made by University staff in interviews, were compared with their degree marks one and two years later. The interview judgments were found to have zero predictive value ; the entrance examination gave rather poor predictions ; but some of the psychological tests gave quite useful indications of later success or failure.

Such results must, however, be interpreted with caution, for two technical reasons. First, the accepted candidates who are followed up may have been actually selected on the basis of the interview. In such a selected group the statistical validity of the selecting instrument

necessarily drops, whereas the validities of tests, which were used only indirectly or not at all, are less affected. Secondly, the investigator usually contrasts the interview predictions with the best of a number of tests, or with the best combination of a battery of tests. Such validity coefficients are by no means stable, especially if the group is small. Different tests might rise to the top on another occasion, and the original battery might show much poorer validity—a phenomenon known as shrinkage. Allowance can be made for these snags, but they are often forgotten by those who condemn the interview as worthless.

A recent, very thorough, investigation which deserves mention is that by Kelley and Fiske, in America, into the selection of students for training as clinical psychologists. The candidates underwent several days of testing and interviewing, and their performance during subsequent training was carefully assessed over four years.

TABLE I. Average Correlations between Procedures applied in selecting Clinical Psychologists and several Assessments of Proficiency after two and after four years (Kelley & Fiske). N = 76 to 93.

Paper qualifications only, judged by two psychologists	0·17	0·22
Paper qualifications + 1 hour's interview	0·15	0·25
Judgments based on paper qualifications + results of objective tests	0·27	0·28
Separate objective tests only : Miller Analogies	0·17	0·30
Guilford-Martin personality inventory	0·22	0·16
Strong interest blank, scored for Psychologist	0·25	0·20
Judgments based on above data + a series of projection tests	0·29	0·26
Separate projection tests, each given by an independent tester.		
Rorschach inkblots	0·12	0·05
Thematic Apperception	0·11	0·12
Group Sentence Completion	0·19	0·21
Ditto + a further intensive interview	0·27	0·26
Judgments of a team based on all the above	0·24	0·30
Final predictions, after observing group exercises	0·20	0·33
Group exercises alone judged by three independent testers	0·27	0·22

The best predictions were obtained from a study of their paper qualifications, and their scores on objective tests of abilities and interests. Group exercises of the War Office or Civil Service Selection Board type were also useful. But neither lengthy interviews by psychologists, nor the application of so-called projection tests such as Rorschach inkblots or Thematic Apperception, added anything to the accuracy of selection. Indeed these methods in which the psychologists felt most confidence were generally the least successful. It seems that the more intensively they studied each individual candidate, the more biased and rigid became their judgments of his vocational suitability. However these conclusions must not be given too much credence, since they are based on less than a hundred cases.

There is some evidence of a more favourable kind. Our own Civil Service Selection Boards similarly make a prolonged study of candidates for the Administrative Grade and the Foreign Service. Account is taken of academic examination results, past record and testimonials ; several intelligence and other tests are given. Then groups of eight candidates are observed while carrying out activities analogous to those of civil servants—discussing some controversial problem, summarising and expounding a brief, taking the chair at a committee, etc. Each is also interviewed by the chairman of the board and by a trained vocational psychologist. I have compared the gradings given at each stage to 330 selected candidates with assessments of their success in the civil service collected $1\frac{1}{2}$ to 2 years later.

TABLE II. Correlations between various stages of Civil Service Selection Procedure and Proficiency after $1\frac{1}{2}$–2 years. N = 330. (All coefficients corrected for selectivity.)

Entrance examinations, or verbal intelligence tests (average coefficient)	0·22
Observation of discussion among groups of candidates (average of three observers)	0·32
Ditto + observation of committee and other exercises (two observers)	0·44
Consideration of all above evidence + individual interviews (two interviewers)	0·47
Final mark after discussion between three observer-interviewers	0·50
Separate board considers all above evidence and re-interviews	0·58

Here you will see that tests and examinations have very poor predictive value, but that with the accumulation of evidence

from the group exercises, the validities rise. The interviews seem to add very little, but actually the judgments from exercises may to some extent have been helped by interview data, as the interviews did not always come last. Note also that a re-interview by an independent civil service board, containing no psychologists, does improve the final predictions significantly.

The value of interviews is attested, also, by the success of the vocational guidance given by the National Institute of Industrial Psychology. Here each candidate takes a few intelligence and aptitude tests, and school records are consulted, but recommendations regarding suitable types of job are based chiefly on an interview by a psychologist, which systematically covers the candidate's life-history, interests, and other relevant qualities. It is regularly found that of candidates who follow the psychologist's advice over 90 per cent. are satisfied with, and satisfactory in, their jobs several years later, whereas among those who do not follow it only some 50 per cent. are equally successful.

It would seem then that some interviewers are moderately good at assessing personality, but that most people—however confident they may be in their ability to sum up interviewees—are not. Unfortunately there is little systematic information as to how to select and train good interviewers ; it is largely a matter of trial and error. Undoubtedly a great many professional psychologists and psychiatrists, as well as employers and teachers, are thoroughly bad interviewers and judges of personality. The selection interview can do some things efficiently. Thus it is useful for marshalling the evidence regarding a candidate's educational and occupational career. Research in the American army shows too that it can be used for assessing such qualities as social aplomb, fluency, good manner and bearing—since these do manifest themselves in the interview situation. But it is an extremely inefficient technique, either for bringing out more fundamental traits like industriousness, emotional stability and moral character, or for synthesising and weighing up the more objective evidence obtainable from past record and test results. The picture that it gives of a candidate's total personality is more likely to be distorted than balanced. Of course we shall have to continue to use it very widely, both because it is the generally accepted method, and because of the difficulties of providing anything better. But the addition of even a few scientific tests would be tremendously valuable in correcting its bias and subjectivity.

Personality Ratings

Now it might be thought that the interview is unsatisfactory because it entails an artificial and strained situation in which the interviewee, particularly if he is a child, can hardly be expected to act naturally. Would not everyday observations of the person's spontaneous behaviour at home, at school, or elsewhere, be better ? But here too we come up against the fallibility of human judgment. There have been innumerable studies where ratings or assessments of the traits of a group of persons have been made by several acquaintances, often on carefully constructed rating scales. Most of these show that the agreement between different raters is just about as low as it is between different interviewers, namely correlations of around 0·5. Higher coefficients are sometimes obtained, say from two teachers ; but this is probably due to their having developed similar points of view or biases, or to having discussed the children beforehand. An illuminating investigation of the value of teachers' ratings of 11-year pupils was included in McClelland's work on secondary school selection at Dundee. Quite good predictions of success in the secondary school were obtained both from intelligence and attainment tests and ordinary examinations, and from primary school teachers' estimates of ability in English and Arithmetic. But when account was taken of their judgments of industriousness and other relevant personal qualities, there was no improvement, indeed there was a decline, in the accuracy of selection. This does not mean that ratings of industriousness showed no agreement with secondary school work, but rather that whatever the teachers knew about their pupils' personalities was already embodied in their English and Arithmetic assessments ; hence the addition of personality judgments was

of no help. Other researches have indicated that teachers' ratings of such traits as persistence among secondary pupils are less accurate than fellow-pupils' ratings of one another ; also that fellow cadets or recruits are better able to predict officer quality than are the recruits' Commanding Officers.

An outstanding feature of ratings is known as the halo effect. When raters assess several qualities, say intelligence, emotional stability, honesty and sociability, it is always found that their judgments of these qualities overlap or intercorrelate very highly. They fail to distinguish what should be largely independent traits, because they tend to see those they are rating through a kind of halo or general impression, and give high ratings on all desirable qualities to those they like, low ratings to those they dislike, regardless of the supposed meaning of the traits. Moreover this tendency is accentuated by closeness of acquaintanceship. The better we know a person, the more crystallised or set becomes our attitude to him, whether favourable or unfavourable, and the less are we able to assess his behaviour impartially, on its own merits. The interviewer, or the observer of group exercises, thus has a certain advantage in coming fresh, as it were, to each candidate.

The Structure of Personality

Here we must venture for a little into the jungle of personality theory. I hope that we shall emerge better equipped to understand the difficulties of assessing personality either by methods involving subjective judgments or by objective tests. It would, I think, be admitted by most psychologists that every child is born with strong drives or needs which are sometimes gratified and encouraged, sometimes frustrated or opposed, by the adults who look after him, and by the other forces in his physical and social environment. His genetic constitution also includes various factors of physique and health, and characteristics of his nervous system and endocrine glands, which underlie his temperamental traits, for example his placidity or excitability, and his abilities. However, it is doubtful whether we can ever observe this innate temperament directly, since even in earliest infancy it is influenced and conditioned by the parents' or nurses' handling, the feeding and sleeping routine, etc. Through the interactions of this growing organism with its environment, an integrated psychological structure is built up which includes the conscious sentiments, interests and habits (as described by McDougall), and the unconscious ' mechanisms ' or complexes of the psychoanalyst. All the motives and past experiences which determine the child's or adult's behaviour in any situation are parts of this exceedingly complex, yet relatively stable, structure that we call his personality. One may or may not accept the details of psychoanalytic theory, but one must agree that considerable progress in disentangling the fundamental principles of personality dynamics has been achieved, notably through the researches of medical and of experimental depth psychologists. I have not the time to refer to more than three of these principles. First, it is clear that many of the links and ganglia in personality structure are repressed into the unconscious mind. This can be proved experimentally by scientific research such as that of Luria. Hence behaviour is often inexplicable at the conscious, rational level, and this is one reason why we can never reach any certainty in tracing out a person's motives. Secondly, Allport has pointed out that fresh mechanisms and interests continue to develop during a man's lifetime, and to become self-supporting or functionally autonomous. Although I agree with Angyal that personality structure, like a musical composition, is a Gestalt in time, yet I see no justification for the view that the motive forces of behaviour invariably trace back to innate instincts, or to Freudian complexes which are fixed during the pre-school years. Indeed one of the reasons for the poor validity of psychologists' and psychiatrists' interview judgments may be that they concentrate overmuch on studying personality longitudinally and historically. A third principle, arising from the field theory approach of Lewin and others, is that personality structure comprises more than the individual himself. It is intimately linked with his social environment, and capable of considerable restructuring in new environments. Thus it is

possible for a child to display remarkably different personal characteristics at home, in school classes under different teachers, or among his friends, and so forth. Recent research on leadership, for example, has shown that this quality is not, as it were, a fixed property of the individual, but varies according to the kind of group of which he is a member and the activities in which the group is engaged. Of course this is a matter of degree. The layman as well as the applied psychologist normally assumes that each individual does possess consistent and distinct traits, attitudes and interests, in almost any social context. Social life could not go on as it does if this assumption was quite unjustified. But the low correlations between different interviewers or different raters are certainly in part attributable to such variations in the individual himself, in part to vagaries of judgment.

Personality Traits

Now a personality trait or quality is an abstraction from this complex structure. I prefer to regard it as a descriptive term, a classification of certain persistent features of peoples' behaviour, rather than—like Allport—as explanatory. Any particular piece of behaviour in fact depends on such a multiplicity of factors in the personality structure and the environment that no-one reacts in accordance with a trait all the time. Nevertheless it is clear that some people behave more markedly or more frequently than others do in a manner that most of us would call sociable, or honest, and so forth ; and that is what I mean if I say that such people are high in the traits of sociability or honesty. There are, of course, thousands of trait-names for describing personal behaviour. But these overlap so much that we can reasonably hope to classify them under a few major dimensions, by means of factorial and taxonomic researches like those of Cattell and Eysenck.

But we are not at the end of our difficulties. How are we to tell which people are, say, persistently sociable or unsociable? Traits are usually rather vague and ambiguous in meaning ; they are not objective like height and weight, in measuring which any two observers would arrive at practically the same result. Mechanical ability, arithmetical attainment and the like also comprise certain types of behaviour which can be fairly readily recognised and agreed upon by different observers (though even here the notorious unreliability of examinations shows that their objectivity is limited). Sociability, honesty, aggressiveness and so on are still more equivocal. Behaviour which one person interprets as aggressive might be called adventurous by another, or limelight exhibitionism by a third. In fact, assigning a trait to an individual inevitably involves subjective interpretation. A purely factual description of behaviour does not give us the personality ; we have to infer or intuit the motives and feelings behind it, and integrate what we see into a picture of the total personality structure, before we can describe a person's traits. Such subjective interpretation is affected by the viewpoint, the values and other personal characteristics of the observer himself. For example, a Victorian parent observes naughty actions in a child, and attributes them to a trait of rebelliousness, where a progressive parent either notices nothing, or regards similar actions as a sign of the child's initiative. Again a Freudian and a Jungian analyst arrive at very different interpretations of the structure of similar neurotic personalities. We cannot say that one is right, the other wrong ; the personality is partly a function of the interpreter as well as of the person interpreted. Though the layman may not be aware of it, he too, like the psychoanalyst, sees people through the distorting spectacles of his own theories of personality. During his lifetime, his introspections about himself, his observations of others, and the books or newspapers he reads and cinema films or plays that he sees, all help to build up in him a set of conceptions of personality mechanisms, of the content of personality traits, and a number of stereotypes or stock personalities. Each new acquaintance that he meets tends to be fitted into one or other of these pigeon-holes, and behaviour which fails to conform is often not noticed or misinterpreted. A fruitful line of study, which few but Dr. A. F. Watts have begun to explore, is the interplay between our linguistic education and our conceptions of people.

Subjective and Objective Methods of Assessment

This all sounds very complicated, but it is just because it is so complicated that mankind has made so little progress in understanding, assessing and controlling human beings. It is hardly surprising that different interviewers or raters often disagree so markedly. A's impression of B's personality is never based solely on impartial observations of B's behaviour in a wide variety of situations. It is always affected both by A's stereotypes or ideas about people, and by the emotional reactions aroused during the interaction of B's personality with his own. On the one hand, B actually reacts differently and shows a different personality to every A he meets. On the other hand, A is very apt to jump to conclusions on superficial evidence and to over-simplify B's structure. As he gets to know B better, his impression of him certainly develops, but it also becomes more rigid, more bound up with an emotional attitude or sentiment, and B's own behaviour in relation to him becomes more crystallised. Thus when A is asked to give judgments of B's traits, his ratings derive from a whole host of more or less unanalysed recollections and inferences and emotional responses, which have been integrated into this sentiment. He may state that A is very honest, or very unsociable, because of such-and-such behaviour, but these reasons consist largely of what the psychoanalyst would call rationalisations.

The clinical psychologist or psychiatrist is more aware than the layman both of the complexities of personality structure and of the need for caution and suspended judgment in interpreting the personalities he is attempting to diagnose. But he is deluding himself if he thinks that the full case-study which he produces after intensive observation and interview is necessarily more objective than the unsophisticated layman's picture. Often he is more ridden by preconceived stereotypes and theories, and the deeper he penetrates the more he may fail to see the wood for the trees, the more biased he may become. Some psychologists, therefore, such as Dr. Eysenck, conclude that this clinical or ' idiographic ' approach should give way to the psychometric or ' nomothetic ' approach based on objective personality tests. But this too has serious weaknesses. We have seen that behaviour which is symptomatic of personality varies greatly with the situation, and an important part of the situation is the attitude of the subject being tested. Thus we find that tests which work well with one group of subjects often fail to do so with another group which is differently motivated, or which is suspicious of the tester. The psychologist's tests, it should be realised, are not miraculous instruments for revealing things which could not be discovered by any other means. They are merely samples of behaviour obtained under specified conditions (including not only overt actions but also verbal responses and autonomic or other physiological reactions as kinds of behaviour). It is essential therefore to include a wide variety of behaviour samples in any attempt to assess personality or measure personality traits, to discard the samples which are too trivial or too unrepresentative, and to combine only those which can be proved valid, either by correlation with some external criterion (e.g. success in an occupation), or by the internal evidence provided by factor analysis. This can be, and has been, done in a number of researches such as Hartshorne and May's, Eysenck's and Cattell's. But it is not easy, and, as I have already indicated, it is scarcely possible in everyday guidance and selection, as distinct from experimental investigations. Nevertheless such work has been carried to the stage where we do possess a number of useful and relatively objective methods which could be given with a view to supplementing, checking and correcting the vagaries of more subjective methods like the interview.

Objective Personality Tests

Physical and physiological tests are not very promising. Sheldon's somatotypes are worth recording, but their validity as indicators of personality traits is far lower than he claimed. Measures of pulse, blood pressure, and biochemical functions appear to vary tremendously from day to day, and their significance, together with that of the psychogalvanic reflex, is still uncertain. The electroencephalograph too has proved disappointing, except in the

diagnosis of epilepsy. However, the body sway test, as developed by Eysenck, generally works well, at least with adults. The extent to which a person sways when standing with his eyes closed, especially when subjected to repeated suggestions that he is falling, correlates moderately with the trait of emotional instability or neuroticism.

Followers of Professor Spearman have done much work with simple tests of perseveration or rigidity *vs.* flexibility, of variability or oscillation, and of fluency or richness of association (*p*, *o* and *f* factors). Such tests tend to be highly unreliable and specific ; the person who fluctuates a lot at one task is not necessarily variable at others. And they are obviously too trivial, and too dependent on the subject's attitudes to constitute good samples of personal behaviour. C. A. Rogers, a student in my Department, has recently found written fluency tests to depend almost wholly on intelligence and verbal ability ; that is, they are unrelated to personality in normal children. However, measures of fluency of speech did show some connection with extraverted traits. The so-called Will-Temperament tests of June Downey are equally artificial, and I have yet to see any good evidence that objective tests of Level of Aspiration indicate the manner in which a person's self-esteem operates in his real-life behaviour. Small group differences do not prove their value for individual diagnosis.

It is perfectly possible to adapt more realistic or natural situations in studying personality. For example, useful tests of persistence have been based on the time a subject will sit with one foot raised an inch above the seat of another chair, how long he will go on trying at some difficult verbal or manipulative problem, etc. In their investigations of honesty, unselfishness, persistence and self-control, Hartshorne and May devised numerous very ingenious tests of this type, which provided genuine samples of the traits. Incidentally they showed how poorly teachers' ratings of the same traits correlated with the objective measures. Among younger children who can be observed for long periods, say in a nursery school, the so-called time-sampling of recording actual instances of specified types of behaviour yields an even more direct approach to trait measurement. But this could hardly be used for guidance or selection purposes.

Indirect Tests

A fundamental difficulty of objective testing is that anyone who knows that his personality is under test is unlikely to behave very naturally. Thus I agree with Professor Burt that indirect approaches are more promising. If the subject supposes that his ability or skill is being measured, then he may reveal important traits in his manner or reacting. For instance, a child who is taking a Binet or a performance test often displays his persistence, excitability, and so on. The tester must, of course, use his subjective judgment in assessing such traits, and he is liable to jump to conclusions on very slight evidence. However, more objective indices of personality traits can be derived in such tests as Culpin and Smith's dotting machine, Cattell's C.M.S. test, the Porteus Maze Q score, the Rey-Davis learning test, Mira's myokinetic diagnosis, and various 'stress' tests developed in America. Still more useful ones might be constructed. We need some machine which will set the subject an intriguing, but difficult and frustrating task, which will also record in easily measurable form the steadiness and persistence of his attack, his preference for speed or accuracy, the extent of breakdown under stress, and his capacity to recover after breakdown. It may be useful to combine this with measures of internal tension, such as the psychogalvanic reflex or the Luria apparatus. One cannot, of course, guarantee that subjects, particularly adults, will react to such a task in the presence of a psychologist as they would to provoking situations in everyday life ; but there is a good deal of favourable evidence.

A number of other types of spontaneous movement—gestures, voice and style of speech, and handwriting—are to some extent expressive of personality, though they are untrustworthy because of their dependence on subjective interpretation by an observer, or a graphologist. However, certain characteristics of handwriting pressure seem to be diagnostic, and these can be objectively recorded. Various measures of speech have also shown promise, for

example the Verb/Adjective ratio, though the recording and analysis of these is probably too elaborate for everyday use.

Group Observational Methods

This brings us to the group exercises—the 'house-party' method—introduced by War Office Selection Boards, and widely adopted nowadays for selecting higher civil servants, industrial executives, ministers, teachers and youth leaders, and even secondary grammar school pupils. It has the merit of being highly acceptable both to selectors and candidates, and it clearly provides a broader basis for judgment than does the typical selection interview. I have been able to show that the predictions made after two days' testing, observing and interviewing at Civil Service Selection Boards correlate almost as highly with satisfactory performance in the Civil Service two years later as do judgments made by supervisors on the basis of more casual observation of the whole of the first year's work in the service. But it should be pointed out that these successful predictions were made by an unusually experienced and stable staff, hence it does not follow that an *ad hoc* board of observers and interviewers, using any old adaptation of group methods, would be equally good. There is not yet in fact adequate validation of the abbreviated procedures introduced, on grounds of economy, into the Army and Civil Service, nor of the other applications mentioned above. Group exercises are not, of course, objective tests; nothing is measured; hence their value is entirely dependent on the skill and impartiality of the particular observer. One of the alleged advantages is that people are observed interacting in a group, not as isolated individuals. But this makes for much greater complexity and difficulties of interpretation. Moreover, it is likely that an individual's behaviour will vary to some extent according to the nature of the particular group to which he is assigned. Finally, it is obvious that candidates realise they are under test. Even though they are nominally engaged on some practical task such as moving an object over a wall, or working out a problem in committee, each one is doubtless trying to display what he conceives to be a 'good' personality. How far their normal characteristics can be masked by such acting is not known. I would conclude then that group exercises provide a useful extension of interviewing, but that they certainly do not meet our need for scientific methods of personality assessment.

Projection Tests

Much the same conclusion holds for what are known as projection techniques. These have shown a fungus-like growth in recent years. The underlying idea is that an individual expresses his personality in everything he does, but especially in his creative and imaginative activities, in his associations to words or inkblots, in his literary or artistic productions, or in play. As a clinical tool for revealing, as it were, sore spots in personality, Jung's word association method has done useful service. The drawings and paintings, and the free play, of maladjusted children are of value both in diagnosis and treatment. There is ample evidence, too, that Thematic Apperception and Rorschach inkblots can help in the differential diagnosis of abnormal personalities. But these techniques have failed signally to prove their worth in educational and vocational selection and guidance. For example, there is no evidence that the Thematic Apperception pictures used in War Office and Civil Service Selection Boards ever contributed any worthwhile information. Apparently they reveal the facets of personality structure which are of interest to the clinical psychologist, but not the overt trends that determine practical adjustments in daily life. Moreover, they are essentially subjective and dependent on the training, skill and insight of the tester, although fair reliability of scoring and even of interpretation have been demonstrated in the case of the Rorschach. Numerous unsuccessful attempts have been made to render them more objective, so that they might provide measurements of important traits. The Kent-Rosanoff technique of scoring word associations, also the mean association reaction time or psychogalvanic response, have no clear significance. Munroe's inspection technique of scoring group Rorschach looked promising, but has not stood up to later research. Some of the

adaptations of the Pressey Cross-Out technique and Maller Word-Connection test do yield scores which correlate sufficiently with other measures of emotional instability or of immaturity to be worth using. Perhaps the most fruitful of the techniques is the group Sentence Completion test, where the subjects write short phrases associated with such stimuli as :

I want to know . . .
My father . . .
I failed . . .

Though faking is obviously possible, most subjects do reveal a wealth of emotionally toned material ; and a very simple scoring of the ratio of unpleasant to pleasant ideas gives good correlations with maladjustment.

Personality Questionnaires

The most practicable method for large-scale application is, of course, the paper-and-pencil test. In this field there have been very considerable advances, and a number of useful questionnaire tests have been devised for adults of above-average intelligence. Their value among average or dull adults, and among children, is less certain. The main difficulty about developing them in this country is the shortage of psychologists to carry out the vast amount of work that they entail. Direct questions dealing with the types of job, the leisure pursuits, etc., in which people are interested, or with their attitudes to controversial topics and issues, are often used ; and a great deal is known about the techniques of formulating suitable questions, item analysis and test construction. But in addition to the obvious dangers of faking if they are applied in any competitive situation, there is the snag that people interpret the instructions and questions so diversely. Some state that they like far larger numbers of items than others do, some are much more extreme, others more cautious, in their opinions. These so-called ' response-sets ' greatly distort the results of most questionnaires. Personality inventories of the Woodworth, Thurstone and Bernreuter variety, which assess a subject's emotionality or introversion by the numbers of neurotic or other answers that he accepts, are hopelessly misleading. They are far too much affected by the attitude of the subject, his suggestibility, co-operation or suspiciousness. They consist, in essence, of a series of self-ratings, and self-ratings suffer in an exaggerated form from all the weaknesses that I have pointed out in connection with ratings by others. As a general rule therefore it is unwise to take any questionnaire answers at their face value. Scoring should be empirically derived by comparing the responses of clearly defined groups. For example, the Strong Vocational Interest Blank yields scores for interest in Accounting, Architecture, Psychology, etc., not by counting the items which look as if they were related to these careers, but from those items which are in fact ticked by large groups of accountants, psychologists, etc. The subject's interest scores are based on the resemblance of his pattern of likes and dislikes to those of members of any occupation. Similarly the Minnesota Multiphasic Personality Inventory is scored by the items which paranoiacs, hysterics, depressives, etc., actually choose. In the U.S. Army Air Force, much the most successful personality test was the multiple-choice biographical inventory, whose items had been proved to differentiate pilots who qualified from others who failed their training. Some of the better questionnaires also show great ingenuity in disguising their questions, or in introducing checks to defeat the faker or malingerer. But an even more far-reaching device is being exploited nowadays—the forced-choice question. Here is a sample item quoted by L. V. Gordon :

Tick the statement that applies most accurately to yourself, and put a cross by the one that is least like you.

(*a*) A good mixer, socially.
(*b*) Gives in readily to other people's wishes.
(*c*) Thorough in any work undertaken.
(*d*) Worries about possible misfortunes.

Two of the answers, presumably (*a*) and (*c*), are high in popularity, the other two relatively low ; and it has been proved that people just about as often choose (*a*) as (*c*), and just about as often reject (*b*) as (*d*). But, unknown to the subjects, one answer in each of these two pairs has been found to discriminate well between

subjects with desirable and undesirable personality traits, whereas the other differentiates poorly. I do not know which these answers are myself; thus I could not make myself out to have a desirable personality on this item. Nevertheless if I do happen to prefer the right answers, there is good evidence that my score will correlate with an external criterion of personality. No doubt this kind of question is rather irritating to the subjects, but it does work. The main difficulty with empirically-validated questionnaires is that very large numbers of people must be tested if reliable scoring keys are to be built up.

Another useful variety of paper-and-pencil test is the information tests of interests. Short tests of information about everyday mechanical and electrical matters were applied in the Services during the war, and are still used, for selecting recruits for mechanical employments. Doubtless they depend partly on previous training and achievement, but they do also involve interest. They could readily be extended to other types of vocation. Peel and Lambert have devised a combination of preference + information test which can be applied in selecting eleven-year-olds for technical education.

Making Use of Acquaintances' Judgments

Now though I have stressed the defects of ratings, and given ample reasons for their untrustworthiness, it is still true that in most guidance and selection situations one would like to make use of the knowledge that acquaintances possess about a candidate. Carefully collected ratings are, in fact, usually more valid than any single objective personality test, and they have the enormous advantage that they do not depend on the candidates submitting themselves to test, or on their co-operativeness. They do not even take up any of the candidates' time, though instead they demand a lot of paper work from the raters or testimonial writers. Good ratings, however, are generally obtainable only in experimental investigations where all the subjects are known to a few raters, who can be trained and supervised in filling in the rating scales. It is a very different matter in vocational work where judgments about each individual are requested from different sets of acquaintances, who cannot be supervised, or in education where numerous primary schools put up pupils for each grammar school. Largely at the behest of psychologists, most Education Authorities nowadays get primary schools to include personality ratings on their pupil record cards, and these are often consulted at the 11-year selection stage, or passed on to the secondary schools to guide the future teachers. I am afraid I have entirely lost faith in such ratings. Even if the primary teachers could be trained to attach a uniform meaning to the traits that are listed, and to judge from objective observation rather than from general impression, there would still be so much variation in standards between raters in different schools that it is almost impossible to interpret the assessments. Valentine's quota system provides one useful method of bringing the schools' judgments of general suitability to a common standard. But its accuracy is limited, and it cannot help with traits unrelated to intelligence. The forced-choice technique has been found useful in reducing variations between C.O.s' ratings of American army officers; that is, the raters tick the statements that apply to the person they are rating, these statements having been chosen to be of equal popularity or halo-value but of different discriminatory value. I can imagine an extremely useful inventory along these lines for 11-year selection in Britain. It would contain some items filled in by the pupils about their own interests, and about the traits of their fellows, and other items answered by the parents, but would consist mainly of personality descriptions checked by the primary teachers. Each item would be validated against later success or failure at one or more types of grammar and technical schools, so that the inventory would be scored by two or more empirically-derived keys. As usual the trouble would be getting people to do the work of construction and validation. Another promising line of progress is suggested by the Vineland Social Maturity Scale, which is filled in by the educational psychologist on the basis of information supplied by parents or acquaintances. In

default of such relatively objective instruments, I would suggest that primary school judgments could best be expressed in the form of a guided personality sketch, not in ratings at all. This would be written in the teacher's own words, but a guide prepared by the educational psychologist would suggest the kind of behaviour to observe and the main points to cover. Similar sketches written by secondary teachers would be more useful to Youth Employment Officers giving vocational guidance to school-leavers than would personality ratings. We all know the defects of ordinary testimonials ; indeed they provide useful confirmation of my view that it is impossible to reach any certainty about an individual's personality. But they can be improved. For example, in Civil Service selection the testimonials evoked from acquaintances by rather detailed questions have proved of considerable help to the selectors.

Conclusion

To conclude then : the problem of personality assessment is perhaps the most difficult in all psychology, but there is no need to be completely pessimistic. Much of the spadework has been done, so that we are now in a position to conduct fruitful research in the innumerable fields where personality differences are a vital factor. And although very few instruments are ready to hand in this country, it is mainly a matter of time and hard work before the trained educational, vocational or clinical psychologist might be provided with a battery of relatively objective instruments. These could be used in individual diagnosis and in large-scale selection work in a much more scientific manner than observation and interviewing are used at the moment.[1]

References

Full references to the various tests and methods mentioned may be found in the author's book, with the same title, to be published by Methuen in 1953. Some of the main sources are :

(1) Allport, G. W. (1937). *Personality, A Psychological Interpretation.*
(2) Angyal, A. (1941). *Foundations for a Science of Personality.*
(3) Cattell, R. B. (1946). *Description and Measurement of Personality.*
(4) Cattell, R. B. (1950). *Personality.*
(5) Eysenck, H. J. (1947). *Dimensions of Personality.*
(6) Eysenck, H. J. (1952). *The Scientific Study of Personality.*
(7) Hartshorne, H. & May, M. A. (1928–30). *Studies in Deceit. Studies in Service and Self-control. Studies in the Organization of Character.*
(8) Jenkins, W. O. (1947). *Psychol. Bulletin,* **44,** 54.
(9) Kelly, E. L. & Fiske, D. W. (1951). *The Prediction of Performance in Clinical Psychology.*
(10) McClelland, W. (1942). *Selection for Secondary Education.*
(11) Stagner, R. (1948). *Psychology of Personality.*
(12) Symonds, P. M. (1931). *Diagnosing Personality and Conduct.*
(13) Vernon, P. E. & Parry, J. B. (1949). *Personnel Selection in the British Forces.*
(14) Vernon, P. E. (1950). *Occup. Psychol.*, **24,** 75.
(15) Watts, A. F. (1944). *The Language and Mental Development of Children.*

[1] Delivered Sept. 4, 1952.

THE ANATOMY OF METABOLISM IN HIGHER PLANTS: TOTAL METABOLISM, METABOLIC PROCESSES AND PARTIAL REACTIONS

Address by

PROF. MEIRION THOMAS, F.R.S.

PRESIDENT OF SECTION K

WHAT is it made of? How has it become what it is? What is happening in it now? How will past and present events influence the future of the individual and the race to which it belongs? Asked of living things, these and similar questions invite anatomical thought in terms of physicochemical as well as biological concepts. Botanists define the developmental problems and with increasing precision as the edge of analytical methods sharpens. Physicists and chemists become indispensable later when the analysed parts have been sorted out. The synthetic attack on the defined problem can then be launched.

The axiom is that transformations of matter and energy lie at the root of all biological events. Specific to every organism, or a component cellular part, is its total metabolism. The physiologist accepts this concept, and abstracts from it metabolic processes for special study. Categories are very mixed in the chapter headings: Photosynthesis; Respiration; Acid, Nitrogen or Fat Metabolism, and so forth. Respiration is a traditional term for the absorption of oxygen and a related production of carbon dioxide by a living organ. The gaseous exchange may be an external manifestation of an integral internal process, having a definite beginning and end. Current researches (p. 235) bear on questions concerning the beginning of the process and the integration of events. The outcome of these researches will determine whether or not the traditional term will survive as something more than a counter with which to reckon. Photosynthesis as a term is in a similar case. When analysis and synthesis can proceed no further, the term may still retain a biological meaning. The crucial point may turn out to be that action is or is not attributable to organisation (thanks be for deliverance by an esoteric word!) within a visible or invisible cell structure. The known facts about certain other processes (e.g. the specialised phenomenon known as crassulacean acid metabolism) point to the possibility of direction by organised enzyme complexes. However, the abstractions from total metabolism implied in many chapter headings in current use are plainly made for classificatory convenience. They are artificial and not natural.

In the middle of the twentieth century we must admit that analysis of total metabolism into component fractional processes that still retain their characteristic pattern has not proceeded very far. Worthy progress may, however, be claimed for the study of the ultimate fragments of which abstracted metabolic processes and, consequently, total metabolism are composed. These fragments ('parts in one stupendous whole') are the partial reactions, the physicochemical realities. With the achievement of a correct synthesis from biochemical knowledge about partial reactions the pattern of a metabolic process would emerge. Ideally, analysis and synthesis would meet in the abstracted process, and we would see in the whole the harmonious working of the parts, relating functional actions, whether single events or blocks of events in a process, to the maintenance of individuals or the perpetuation of a race.

Ponderous tones may be heard again in the arguments, reasoned, I hope, but not profound, which I shall develop in chapters on total and intermediate metabolism. Then I shall come down to brass tacks. What we require are well authenticated facts placed in their setting. I shall outline some of the methods used in fact finding, pointing out limitations in the value of evidence so gained. Finally, for three aspects of carbohydrate catabolism in higher plants I shall sketch for each abstracted process what appears at present to be the emerging pattern. Much is still conjectural, especially about higher plants. On each fabric there are blurred patches between clearly marked threads. I shall endeavour, by brief reference to a limited amount of modern work, to indicate what is being done to trace forms in the patches and to bring the threads together. I know that my accounts are not quite up to date. I shall dare to report unpublished work (Newcastle provenance, 1930's), and to make more of my own interests than of those of others.

To observations which ourselves we make,
We grow more partial for th' observer's sake.

The abrupt ending of this Address will symbolise my firm conviction that nothing is rounded off yet.

The Total Metabolism of a Growing Plant

Water culture or sand culture experiments can determine the minimum external requirements in matter and energy for plant growth. Chemical analysis, coupled with knowledge of anatomy and cytology, resolves the nature of the metabolic products present at successive times in the nucleus, cytoplasmic structures, cell walls and vacuoles of the tissues composing a growing plant. The exact chemical description of a plant at a time T demands not only the identification of the molecules present, but also the localisation of these molecules.

A complete statement about total metabolism between two times T_1 and T_2 would require knowledge of changes in the chemical composition of all the distinct parts to which morphologists, anatomists and cytologists have given names, and also of the related changes in the composition of the environment.

The metabolic products have been classified in various ways. Interest may attach to their possible functions. Thus enzymic proteins and some non-enzymic proteins, lipoids, sterols, nucleic acid and all the coenzymes produced between T_1 and T_2 are components of the protoplasmic machinery. Cell-wall substances form a structural group. Food substances immediately available or as storage products, hormones, and other substances which contribute to internal economy may be regarded as of physiological significance. Substances affecting external relations can be grouped together as of ecological significance. Thus nectar substances, pigments and perfumes may promote insect pollination. A dermal covering of wax might ensure the survival of a plant under arid conditions.

A large number of cell components are intermediate anabolites and catabolites, some of which may accumulate only in traces. Some accumulating substances may have no functional value. If they are not lethal the cell would survive the side reactions by which they are produced.

Even when a metabolic product has a clear functional value at T_2 we do not attribute its ontogenetic production to a purposeful cause, but to the interaction of the molecules present at T_1, and of those absorbed between T_1 and T_2, with the molecules of a metabolising system bearing a certain racial inheritance. However, we could argue that the chances of survival of a certain inherited pattern are enhanced commensurably with the number of partial reactions which have a favourable influence during ontogeny. The arguments could be extended to the point of attributing morphological adaptations to environment to favourable qualitative and quantitative changes brought about in partial reactions by mutations in the inherited metabolising systems. These great matters are too high for me.

We are on solid ground in making a classification of metabolic products on a chemical basis. The elucidation of metabolic processes is dependent on knowledge gained by organic chemists of the mole-

cular structure of the substances occupying a certain position and performing a certain function, or doing no harm, in a plant at a given time. Names such as starch, cellulose, lignin, albumen, oil, essential oil, chlorophyll, had been given to plant products long before their chemical structure was investigated. Chemical resemblances and differences are not self-evident, but, thanks to the brilliant achievements of organic chemists, we can now put forward sensible hypotheses about possible sequences in the formation of metabolic products. We no longer attribute the appearance of autumn coloration in leaves to the conversion of a green into a red substance. In its molecular structure chlorophyll, as a magnesium porphyrin derivative, is utterly different from the glycosides of phenyl γ-benzopyryliums, i.e. anthocyanins as a class. Nor do we accept the authority of the *Shorter Oxford English Dictionary* (1933 Edition) that etiolin is a yellow modification of chlorophyll found in plants growing in the dark. Positive assertions can now be made abundantly. We can maintain that, since starch and cellulose are carbohydrates there are some features in common about the anabolic reactions by which a food on the one hand and a cell-wall substance on the other are produced by plants, but quite a different set of reactions must be involved when lignin (a benzene derivative) is laid down in cell walls. The argument can be developed along obvious lines. There are common types of reactions in the formation of all proteins, whether protoplasmic or food, and of all fatty oils, whether physiologically important as food reserves or serving an ecological function in a feeding tissue such as the flesh of an olive fruit. Among the essential oils are terpenes which, as isoprenoid compounds, can be classified with substances of very different functions, the carotinoids, phytol, sterols, and sapogenins. Conceivably all of these isoprenoid compounds could be produced by condensation of C_2 and C_3 fragments of glycolytic cleavage of sugars (see legend to fig. 1), by different combinations of reactions having certain features in common.

There is little doubt that eventually such metabolic powers of living cells as are displayed in partial reactions will all be described in terms precisely defined by chemists.

I have heard it said by an eminent chemist that the mechanism of metabolic sequences will not be really unravelled until each partial reaction in a sequence is carried out *in vitro* with a purified enzymic protein and its substrate, and such coenzymes and co-factors as may be required, as the sole components of the reaction mixture. This statement is true for each partial reaction considered in isolation; but the elucidation of total metabolism in the cell at any time will require the ravelling anew of the partial reactions into the pattern inherent in the cell at that time. The reactivity of the whole cell is the resultant of the integrated activities of its enzymic components in a manner characteristic of the race to which the plant belongs. For any cell the reactions would have to be assembled collectively in a specific way to reconstitute the metabolising whole. I think we must admit that such a physical reconstitution is beyond our present powers. In less austere times I might have demonstrated this point of view by smashing an egg and reciting 'Humpty-Dumpty.' Mental reconstitution is our present aim. Permutations and combinations of similar metabolic powers would lead to qualitative as well as quantitative differences in total metabolism. There are different kinds of eggs.

As knowledge increases the border line between the disciplines of organic chemistry and biochemistry becomes less well-marked; but it is still a defensible assertion that all reactions requiring as one of the reactants a cell substance not yet synthesised by the chemist should be regarded as still outside the domain of organic chemistry. Enzymic proteins in existing organisms, evolved through the ages, are metabolic products formed during ontogeny in each succeeding generation, generally as fit to survive in the struggle for existence, certainly as not unfit, in the balance of chemical events under natural conditions. Chemists move in the direction of synthesising substances that in point of structure and size of molecule merit the name of protein; but they have not yet synthesised a native protein of any kind, enzymic or

non-enzymic. Were they to succeed in doing so, it would be their greatest triumph. This is saying a great deal, for their recent achievements have been vast indeed. For the present we may regard the study of enzymic reactions *in vitro* as biochemical rather than purely chemical.

We should remember, however, that products of certain enzymic reactions may, *in vivo* or *in vitro*, undergo transformations that are independent of the presence of enzymes, but I can think of no spontaneous chemical change in a cell which does not occur at the expense of a metabolic product. All substances found in a plant at a time T are products of evo*lution*ary processes (whatever these may have been) since those remote ages when such ancestral forms of present-day partial reactions as have been pictured by Professor J. D. Bernal (1951) were acquiring potential organismic significance.

Statement of Problems of Intermediate Metabolism

For the complete description in chemical terms of enzymic reactions and such non-enzymic reactions as may occur between two times T_1 and T_2 there are certain obvious requirements. We must seek (*a*) to ascertain what intermediate metabolites accumulate and what transitory substances are produced when known initial metabolites (A_1, A_2, A_3, etc.) are converted into arbitrarily defined anabolic and catabolic end-products (Z_1, Z_2, Z_3, etc.), (*b*) to place these intermediate metabolites (B's to Y's inclusive) in their proper sequence, (*c*) to discern the types of chemical reaction by which protoplasm effects each partial reaction in the sequence, and (*d*) to separate from protoplasm the enzymic systems (e_a, e_b, e_c, etc.) specific for each reaction, to study their properties *in vitro*, and to obtain evidence of possible linkages between different systems.

If we choose, we may regard the A class of substances as the carbon dioxide, water and mineral salts from which are built proteins, carbohydrates, fats, aromatic compounds, isoprenoid substances, or any class of Z anabolic product on which our attention may be fixed. Alternatively our abstraction from the total metabolism between the times T_1 and T_2 may isolate as a unit metabolic process for special study, the catabolism of an anabolic product (this time the A class) with the formation of Z class catabolic products.

If we use a single letter to denote either a single substance or two or more substances we can represent a conversion of A to Z by a straight sequence of successive partial reactions thus :

$$A \underset{v_a}{\overset{e_a}{\rightleftarrows}} B \underset{v_b}{\overset{e_b}{\rightleftarrows}} C \rightleftarrows \ldots Z$$

A number of these reactions may be reversible, but I shall assume a progressive consumption in the direction of Z. Apart from certain short-distance catabolic changes, there are probably not many independent sequences of this kind. In the first place, any one of the intermediate metabolites may have more than one fate, leading to end-products Z_1, Z_2, etc. For example glucose, as an equilibrated mixture of α and β forms, could give rise anabolically to starch and cellulose and catabolically to carbon dioxide and water.

Secondly, continued metabolism in certain directions is dependent on linkages between sequences of partial reactions. Apart from light reactions, anabolic sequences must link with catabolic sequences. This broad statement really means that endergonic partial reactions are driven by the energy derived from exergonic reactions. The synthesis of Y from X depends upon linkage with an exergonic reaction $M \rightarrow N$ such as may occur in the respiratory sequence. Instances multiply of the mediation of adenosine triphosphate (ATP) in driving anabolic partial reactions, owing to the high energy potential of its terminal pyrophosphate bond. Clearly, therefore, all such reactions must link with the oxidative reactions that raise this potential. An example is provided by Chibnall's experiment in which he caused rye grass leaves to synthesise glutamine from ammonium glutamate. He attributed the driving force to respiratory energy. Further, there is evidence that glutamine can be synthesised *in vitro*, provided ATP is present as an energy donor. This evidence suggests that the linkage is not between glutamine synthesis and the respiratory process as a whole (cf. p. 235) but with those partial reactions which lead to

the production of active phosphate bonds and to their incorporation in ATP.

Another form of linkage between different reaction sequences, or between distinct partial reactions in a single sequence is found when, in a reaction such as

$$P + Q \rightleftarrows P' + Q'$$

Q is present in very low concentration. Clearly the conversion of P to P′ will soon stop unless Q can be regenerated from Q′, for example thus :

$$Q' + R \rightleftarrows Q + R'.$$

The net result will be

$$P + R \rightleftarrows P' + R'.$$

Examples of Q are coenzyme I (DPN or cozymase cz), coenzyme II (TPN), glutathione (G–S–S–G), the flavin nucleotides (e.g. FAD), cytochrome c (cyt. c), dehydroascorbic acid (DAA), orthoquinones, or other intermediate carriers. Each of these substances can act as a reaction partner with more than one enzymic protein. Accordingly the partial reactions controlled by these proteins may be linked.

In recent times the results of broad physiological studies and the narrower but more penetrative probings of biochemistry have proved mutually illuminating. In the latter, attention is focused on a fraction of the process abstracted for study from the total metabolism of the cell. In the extreme case the single partial reaction is actuated *in vitro* by the isolated specific enzyme. Biochemists have achieved much success during the last twenty years, in isolating enzymes as crystalline proteins in a high degree of purity. Theoretically, were knowledge gained of all the partial reactions in a process, a mental synthesis of the whole process would become possible. Clearly, however, this synthesis must serve to explain the results obtained from broad physiological studies. This requirement is a touchstone. There are many possible pitfalls in assembling whole processes from enzymic reactions. The enzymic reaction may be an artefact. If it has metabolic significance, it may be concerned with some other process. Even if it is part of the process under consideration it might be placed in the wrong position in mentally constructing the whole process.

Some Methods used in Studying Intermediate Metabolism

The question is whether the change from A to C occurs *in vivo* by way of B. I shall assume that on chemical grounds this sequence is a clear possibility, and that it would be in accord with such facts as have been obtained by broad physiological experiments. For my present purpose it is immaterial whether the conversion of A to C is a two-step reaction in which a single e_a occasions the change A to B with a velocity v_a, and another enzyme e_b transforms B into C with a velocity v_b, or whether in each stage in the conversion there are two or more steps, i.e. both e_a and e_b represent a collection of enzymes, with possibly coenzymes and co-factors in addition.

(i) *Qualitative analysis* (see also vi). The object is to detect the presence of B in cells which are known to have been converting A to C over a period T_1 to T_2. Since B may originate during the course of some other reaction, the evidence, if positive, would be stronger were it found that B is present in cells only under conditions which permit the conversion of A into C. The value of this method for trace accumulants (and many key intermediates belong to this class) has recently been enhanced by the refinement of chromatographic and other techniques. But in many reaction sequences B still eludes detection because, under natural conditions *in vivo*, v_b is so much greater than v_a.

(ii) *Fixation methods.* A substance X, which can enter into chemical combination with B, is added to cells which are converting A into C. The argument is that X will compete for B with the enzyme e_b, and, in consequence, some B may become fixed as BX. Obviously, for the reasons given in (i), positive evidence is not conclusive. Furthermore, X might have a disturbing influence on the normal balance of linked reactions in ordered metabolism, resulting in the formation of B, and therefore BX, as an artefact. A famous example of this method is Neuberg's fixation of acetaldehyde as its bisulphite when he added sulphites to cells which were carrying out alcoholic fermentation.

(iii) *The differential inhibition or activation*

of enzymes in living cells. The object is, by the addition of appropriate chemicals, differentially to inhibit e_b or to activate e_a so that v_a becomes greater than v_b. Obviously B would then tend to accumulate and method (i) should give positive results. A positive result does not provide conclusive evidence because, in this method as in (ii), abnormal chemical changes may occur owing to the presence of the inhibitor. However, this is now one of the most widely used methods.

Presumably all inhibitors combine chemically with or occupy the active surfaces of the enzymes which actuate the blocked partial reactions. Competitive inhibitors have a molecular structure similar to that of substrate on which the enzyme normally acts, and may be specific for a single reaction. For example malonate inhibits the oxidation of succinate to fumarate by succinic dehydrogenase. A depression caused by malonate of the oxygen-uptake of a tissue is considered as evidence that succinic dehydrogenase possibly functions in respiratory metabolism. If succinate accumulates simultaneously with this depression, the evidence is stronger.

Terminal oxidases containing copper (e.g. polyphenol oxidase and ascorbic oxidase) or iron (e.g. cytochrome oxidase) are poisoned by hydrogen cyanide. One of the tests for cytochrome oxidase is to show the presence of a terminal oxidase (i.e. an enzyme concerned with oxygen uptake) that can combine with carbon monoxide to form a compound which is dissociated by wave-lengths of light absorbed by this CO-hæmochromogen. The biochemical results of these purely chemical phenomena are that carbon monoxide inhibits cytochrome oxidase, but the inhibition may be removed by light.

I reported to this section at Bristol in 1930 that ethyl alcohol and acetaldehyde accumulate in apples treated in air with HCN or H_2S. This finding fulfilled predictions based on Meyerhof's experimental results for yeast and Warburg's for animal tissues. I described the phenomena as HCN zymasis and H_2S zymasis, and drew the general conclusions, (*a*) that oxidation processes link with zymase cleavage, and (*b*) that, if a substance such as HCN retards aerobic oxidative processes more than zymase cleavage, zymasis will occur, provided the organ under experiment can carry out this process in the absence of oxygen. My view was, and still is, that it does not matter whether, in such an organ, oxidative processes are retarded by a shortage in the external oxygen supply or by differentially reducing the activity of an enzyme that controls any one of the partial reactions by which C_3 or C_2 products of glycolysis are consumed in air, the result will inevitably be some form of zymasis. As knowledge grows of partial reactions in phosphorylative glycolysis and the tricarboxylic acid cycle, so also does the scope of experiments on the induction of zymasis.

It was in 1924 at Cambridge that I first induced zymasis in a plant organ plentifully supplied with oxygen. The inducing agent was carbon dioxide in concentrations above 30 per cent., and the organ was the apple. The results of this investigation were unpredictable. A shot at random hit a mark. I described the resulting disordered metabolism as CO_2 zymasis. This investigation had a bearing on storage problems. In CO_2 zymasis, in addition to ethyl alcohol, injurious concentrations of acetaldehyde may accumulate and lead to the browning of flesh tissue. [Jacobsen (1952) comments on the nausea felt by man when acetaldehyde accumulates in his liver. 'Aldehyde poisoning' is not confined to apple flesh tissue.] One essential requirement was established for keeping fruit in sound condition in home stores and ships' holds. The storage atmosphere must be so controlled that there should never arise, as a result of the respiration of the stored fruit, such conjunctions of CO_2 and oxygen concentration as might induce CO_2 zymasis.

Newcastle experiments performed successively by Fidler and Foster have shown that HCN zymasis, H_2S zymasis, and CO_2 zymasis may be induced in a wide range of organs. Beevers' results suggest that CO_2 in high concentrations does not belong to the same class as HCN and H_2S, namely the inhibitors of terminal oxidases. We are still trying to identify the partial reaction which CO_2 differentially inhibits. Success would provide evidence

that this reaction might participate in the normal aerobic oxidative consumption of carbohydrate.

The converse of the same principle holds. The induction of zymasis by a substance known to inhibit a putative partial reaction in a hypothetical scheme for such consumption could be a cogent piece of evidence in favour of the scheme. The argument requires careful wording to ensure that the question is not begged.

(iv) *The demonstration of the presence of specific enzymes in living cells.* Enzymes are the ultimate subcellular agents of each action in metabolic activity. If a reaction does not occur spontaneously *in vitro*, evidence of the existence of an actuating enzyme is needed before biochemists will agree that the reaction is a possible step in metabolism. The demonstration of the presence in living cells of e_a and e_b clearly proves that the cells are potentially capable of converting A to C *via* B, but does not establish that this is the route by which the conversion actually occurs. Owing to organisation within the structure of a living cell, a substrate may not always have access to an enzyme which can act on it.

If a thermolabile cell-free product, whether an aqueous extract or an insoluble residue, can bring about a change known or strongly believed to take place *in vivo*, acceptable evidence of the presence of an enzyme is obtained. Artefacts are possible in the form of changes which have no metabolic significance but happen to be catalysed by certain cell proteins. Many enzymic proteins have been highly purified, but very few from higher-plant extracts. Some coenzymes have been synthesised.

(v) *Feeding experiments.* If B is fed to a cell capable of converting A to C, and it is found that B is consumed and the production of C is commensurably increased, evidence consistent with the hypothesis that B is an intermediate in the conversion $A \rightarrow C$ will have been obtained. Alternatively, exceptional conditions allowing the accumulation of B, for example the exclusion of oxygen, can sometimes be arranged. The fate of B can be studied when normal conditions are re-imposed. There could be other explanations, for instance, that D and not B is the real intermediate, but there is an enzyme present that can convert B into D. Suggestive results have already been obtained by feeding plant tissues with compounds containing labelled atoms. I shall leave it to a future President of this Section to deal with those conclusions which weather the years.

Feeding experiments have been used in investigating the status of ethyl alcohol and acetaldehyde as possible normal intermediates in the oxidative catabolism of higher plants. At Cambridge I caused these substances to accumulate in apples, by placing the fruits under anaerobic conditions, but found little consumption on replacing the fruits in air. This negative result for respiring tissue was obtained late in the 1923–24 storage season. At Newcastle, Fidler, Foster (also at Durham) and Ranson have successively obtained positive results on a wide range of organs. Our general conclusion is that ethyl alcohol can be slowly consumed in air, and acetaldehyde more rapidly. Remembering my negative result, I have selected an auxiliary verb of mood to state a particular proposition.

Owing to its low boiling-point, acetaldehyde can be fed as a vapour to plant organs. At Cambridge, Fidler (1937) has repeated and extended earlier experiments by Kidd and Trout. Dr. Fidler has kindly allowed me to read a fuller report of his experiments and to refer here to his results. They show that living orange and apple fruits can absorb and consume acetaldehyde extensively. The absorbed acetaldehyde does not accumulate : to some extent it is converted into ethyl alcohol. Fidler measured gaseous exchange, carbohydrate loss, as well as alcohol production, and from his balance sheet he educes strong evidence that most of the acetaldehyde which is not reduced is completely oxidised to carbon dioxide. Foster's direct feeding experiments, using a wide range of organs, have also shown aerobic consumption of acetaldehyde, only part of which can be attributed to reduction to ethyl alcohol. The aerobic consumption of ethyl alcohol was always too slow for this substance to be considered as a possible normal intermediate metabolite in the aerobic oxidation of carbohydrate.

The question remains open whether or

not acetaldehyde is a normal intermediate in this process. The results of the experiments performed at the Low Temperature Research Station, Cambridge, strongly suggest that it can be easily oxidised by apples and oranges. Indeed Fidler infers with good reason that it can spare carbohydrate. However, it is possible that, like ethyl alcohol in the human liver, acetaldehyde (B of the opening paragraph of this section) is oxidised in plant cells because they happen to contain an enzyme which can convert it into a metabolite (D), normally produced by partial reactions not involving carboxylase decomposition of pyruvate, i.e. acetaldehyde formation. Among possible D substances are acetate and acetylphosphate. They could arise from pyruvate, without acetaldehyde production, by mechanisms similar to those in animal cells. At present nothing is settled for higher plants about partial reactions in oxidative catabolism at the pyruvate and C_2 levels.

(vi) *The accumulation sequence.* Qualitative analysis may show that B accumulates before C when A is being converted into C. When these relations hold, accumulation would provide evidence that B is an intermediate in this conversion, but not rigid proof. The sequence might be $A \xrightarrow{v_a} C \underset{\leftarrow}{\overset{v_b}{\rightleftarrows}} B$, with a velocity v_b so much greater than v_a that the accumulation of B would come first, and that of C not until the back reaction sets in. The order of accumulation is not necessarily that of production.

(vii) *Privative methods.* An essential factor operating in the reacting system A may be excluded. This deprivation would result in stopping the conversion of A to C, but that of A to B might continue. It can be argued that the accumulation under these conditions of B, or of a compound B′ clearly formed from B, provides evidence that B is an intermediate between A and C under normal conditions. However, B or B′ might have been derived from D.

An application of this method is the comparison of the anaerobic and aerobic respiration in higher plants. Ethyl alcohol, which accumulates when oxygen is excluded, is regarded as a B′ substance. We do not yet know for certain what is the precursor, B, common to ethyl alcohol and the products of the complete oxidation of carbohydrate, viz. CO_2 and water. A good case can be made for phosphopyruvate, or a derivative.

(viii) *The value of cumulative circumstantial evidence.* None of the methods described in this section can provide data for the formulation of categorical propositions about the occurrence and sequence of partial reactions in total metabolism *in vivo*. All too frequently statements are made in critical arguments that ' X has shown so and so,' when X's actual contribution has been to obtain circumstantial evidence, sometimes by only one experimental method. X may have stressed a preference for one interpretation of the evidence. He has a right to do so. Clearly, however, the hypothetical rather than the categorical proposition should, as a rule, be used in the incorporation of X's findings and conclusions into one's own arguments.

Granted that all evidence relating to intermediate metabolism is circumstantial, yet as Sherlock Holmes remarks in *The Noble Bachelor* ' Circumstantial evidence is occasionally very convincing, as when you find a trout in the milk, to quote Thoreau's example.' Cumulative circumstantial evidence increases one's confidence in the validity of the conclusions drawn. Prosecuting counsel would make good play with the presence in the milk of a little river sand and of freshwater algae, vouched for by an expert witness from Section K.

Carbon Dioxide Fixation as a Key Reaction in Crassulacean Acid Metabolism

In crassulacean acid metabolism, malic acid accumulates at the expense of carbohydrate in green cells after they have been transferred from the light to the dark. This dark acidification continues for some hours. Whilst this process is operating, CO_2 *liberation* from the cells is less than oxygen uptake. Indeed the R.Q. can approach or reach zero. During the course of dark acidification the oxygen uptake remains steady, but, eventually, the CO_2 liberation increases, and, in consequence, the R.Q. rises and reaches a value of about unity when malic acid

ceases to accumulate. In prolonged darkness the malic acid concentration falls. The R.Q. rises, and, during this dark deacidification, may reach values of two or more. Accordingly it is a widely accepted theory that there are three phases in crassulacean acid metabolism in the dark. In the first the rate of malic acid production is faster than that of consumption. In the second, the rates are equal (Wolf's 'stationary phase'), and, in the third, consumption is faster than production.

When the green cells are illuminated consumption becomes so much faster than production that light deacidification can cause most of the malic acid produced overnight to disappear during the day. The assimilatory quotient (O_2 evolved/CO_2 absorbed) can be much higher than unity: indeed there may be oxygen evolution without CO_2 absorption.

In the Wood-Werkman reaction, pyruvic acid or a derivative combines with carbon dioxide to yield oxaloacetic acid, which is then reduced to malic acid (see fig. 1). [Gollub and Vennesland (1947) have reported that this CO_2 fixation was actuated by a β-carboxylase, which they extracted in cell-free juices from parsley roots and other plant organs.] In 1945 I saw, and in 1947 I reported, that if a Wood-Werkman reaction can occur in green cells showing crassulacean acid metabolism, known facts could be explained.

In the dark, respiration would provide a source of CO_2 for malic acid synthesis.

$$C_6H_{12}O_6 + 6O_2 = 6CO_2 + 6H_2O \quad . \quad . \quad (i).$$

This synthesis would take place at the expense of pyruvic acid (or a derivative), formed by phosphorylative glycolysis (fig. 1). The overall equation for the production of acid from sugar would be

$$C_6H_{12}O_6 + 2CO_2 = 2C_4H_6O_5 \quad . \quad . \quad (ii).$$

During dark deacidification only a fraction of the CO_2 *produced* by respiration would be *liberated*. An explanation would be afforded of the inverse relation between the rate of acid synthesis in the dark and CO_2 liberation. The *apparent* R.Q. would be less than unity. If all the respiratory CO_2 of process (i) were fixed by process (ii) the apparent R.Q. would be zero.

$$2C_6H_{12}O_6 + 3O_2 = 3C_4H_6O_5 + 3H_2O \quad . \quad . \quad . \quad . \quad (iii).$$

Equation (iii) suggests the old hypothesis that malic acid accumulation relates to a single respiratory process in which carbohydrate is incompletely oxidised. I discarded this hypothesis. Equation (iii) is the resultant of the two linked processes (i) and (ii). In the first carbohydrate is completely oxidised in normal respiration. In the second malic acid is synthesised by partial reactions that are not respiratory. Carbon dioxide is not evolved; it is absorbed. Actually, all the time the real R.Q. is unity, but it can only be measured during the stationary phase. The high R.Q.'s during dark deacidification appeared to be the best explained on Bennet-Clark's hypothesis that at least some of the malic acid is reconverted into carbohydrate.

$$2C_4H_6O_5 = C_6H_{12}O_6 + 2CO_2 \quad . \quad (iv).$$

The possibility of consumption in other ways is under consideration.

In the light, respiratory CO_2 as well as the CO_2 absorbed from the atmosphere is immediately removed by photosynthesis.

$$6\,CO_2 + 6H_2O = C_6H_{12}O_6 + 6O_2 \quad . \quad (v).$$

Acid production (equation (ii)) stops; but acid consumption (equation (iv)) continues, and the CO_2 produced spares atmospheric CO_2. Consequently the assimilatory quotient rises above unity.

At Newcastle in 1946 Beevers performed experiments on leaves of *Bryophyllum* and *Kalanchoë* to test some of these ideas. He showed that by augmenting the CO_2 concentration around illuminated leaves which have initially a low acid value, he could induce light acidification (equation (ii)). He had supplied the missing metabolite, viz. CO_2, for the Wood-Werkman reaction. Even more convincing were the results of experiments in which leaves were fed in the dark with CO_2 in concentration up to 5 per cent. During the initial burst of dark acidification, oxygen and carbon dioxide were simultaneously absorbed. Not only was all the respiratory CO_2 (equation (i)) consumed in malic acid synthesis (equation (ii)), but additional CO_2 was required to satisfy the acid-synthesising powers of the cells. Ranson has extended the 1946 work in

many directions. In every respect his results have substantiated the statements I made at a meeting of the Society for Experimental Biologists in London in January 1947, and in a book later in that year.

In the U.S.A., Bonner and Bonner (1948) reported that ' The formation of organic acids by excised leaves of *Bryophyllum crenatum* in the dark at low temperature is approximately a linear function of CO_2 pressure in the surrounding atmosphere between the levels of 0·00 and 0·1 per cent. of CO_2 in air.' This and certain other experimental findings led them to suggest that ' the formation of acids by succulents results from carbon dioxide fixation, perhaps by pyruvate, with the formation of four-carbon dicarboxylic acids.' Independent researches at Newcastle and Pasadena, performed at about the same time, had led to very similar conclusions about the central importance of the Wood-Werkman reaction in malic acid synthesis, and therefore in the control of crassulacean metabolism. Thurlow and Bonner (1948), using $C^{14}O_2$ demonstrated that as a result of dark fixation of CO_2 by Bryophyllum leaves, labelled carbon becomes incorporated in the accumulating organic acids. The metabolic consumption of CO_2 in a partial reaction leading to malic acid synthesis was proved beyond reasonable doubt. The mechanism of this synthesis is still under investigation. A single protein, Ochoa's ' malic enzyme ' may effect the synthesis. Alternatively malic acid may be produced from pyruvic acid and CO_2 in two partial reactions, the Wood-Werkman reaction under the agency of oxaloacetate carboxylase, followed by the enzymic reduction of the resulting oxaloacetate. Vennesland *et al.* (1949) have obtained evidence of the presence of the malic enzyme as well as of oxaloacetate carboxylase in extracted plant juices. The alternative modes of synthesis are pictured in fig. 1.

I shall end this chapter with a historical note. Theodore de Saussure (1804) recorded the simultaneous absorption of CO_2 and oxygen in the dark by the cactus *Opuntia*. The few physiologists who noticed this observation attributed the CO_2 absorption to solution in cell sap. Before the end of last century it was known that cacti are among the plants which show crassulacean acid metabolism. It is a fact of psychological interest that no one attempted to relate the dark absorption of CO_2 to dark acidification. The absorption had to be explained away as a physical and not a metabolic phenomenon. The minds of physiologists were so conditioned that metabolic consumption of CO_2 in green plants, except by illuminated chloroplasts, was inconceivable. Actually the significance of Saussure's observation was not realised until after the recent Newcastle and Pasadena experiments had shown that CO_2 fixation accompanies the dark accumulation of malic acid. In 1949 Wolf, in Germany, and I paid tribute to the great experimentalist who, over 140 years earlier, had recorded the first evidence of dark CO_2 fixation. Independently we referred to this dark CO_2 absorption by plants showing crassulacean acid metabolism as the ' de Saussure effect.' At Newcastle, Ranson has found that in all such plants as he has examined, including *Opuntia* and other cacti, the effect is shown, and to an extent commensurate with malic acid accumulation.

Anaerobic Respiration

(i) *The term zymasis.* In 1926 I introduced the term, *zymasis* to describe the process of alcoholic fermentation, abstracted from the *total* catabolism of a cell. The essential property is that ethyl alcohol accumulates at the expense of hexose units. The word fermentation, when unqualified is ambiguous. For example lactic acid can be produced by non-zymasic fermentative processes. Further the word zymasis allows pithy compound terms to be coined indicating both the occurrence of alcoholic fermentation and the occasioning cause. Among the forms of zymasis I have so far described are anaerobic zymasis, low oxygen zymasis, CO_2 zymasis (p. 224), HCN zymasis (p. 224), and, senescence zymasis (e.g. zymasis shown by apples and pears in air, owing possibly to depressed oxidative power in their old age).

(ii) *Anaerobic zymasis.* In the organs of some higher plants (e.g. carrot roots, germinating peas) the CO_2 of anaerobic

respiration (I) comes almost exclusively from zymasis according to the equation

$$C_6H_{12}O_6 = 2CO_2 + 2C_2H_5OH.$$

This finding provides assured evidence that the partial reactions of zymase-cleavage are integrated into what may be regarded as a single anaerobic metabolic process. If Z represents ethyl alcohol accumulation plus the small accompanying amounts of acetaldehyde, and Z_{CO_2} the corresponding CO_2 production, as a close approximation we may state that $I = Z_{CO_2}$, and in consequence $Z/I = 1{\cdot}04$. For some organs (e.g. resting potato tuber) this quotient is well below 0·5. Like most physiological magnitudes and unlike the physical constants of pure chemicals, the quotient is highly variable. At Newcastle, Fidler found values varying from 0·5 to 0·9 for different varieties of apple and noticed seasonal fluctuations in the quotient for a given variety.

All the available fragments of experimental evidence (see p. 232 for evidence obtained by Stumpf) fit in with the idea that the partial reactions in the zymasis component of anaerobic respiration in the higher plants are identical with those brought about by zymase cleavage of hexoses in yeast cells. Anaerobic phosphorylative glycolysis at least as far as phosphopyruvic acid, appears to be identical with that occurring aerobically (fig. 1). Subsequently pyruvic acid is produced and carboxylase decomposes it to yield Z_{CO_2} and acetaldehyde. It is still unsettled whether or not these two last events are normal partial reactions in aerobic metabolism (p. 226). Almost certainly the next event does not take place during normal respiration in air. Under the agency of alcohol dehydrogenase, acetaldehyde accepts hydrogen from the $DPNH_2$ which is formed in the anaerobic oxidation to phosphoglycerate of triosephosphate by its specific dehydrogenase. This reduction yields ethyl alcohol, and DPN (cozymase) is regenerated. Accordingly the oxidation actuated by triosephosphate dehydrogenase is linked with the reduction of acetaldehyde to ethyl alcohol. Since very little acetaldehyde accumulates, this reduction, which involves the oxidation of $DPNH_2$ to DPN, must be a relatively fast process. The rate of turn-over of DPN $\rightleftarrows$ $DPNH_2$ may control the rate of triosephosphate oxidation, and therefore of glycolysis.

The problem of determining the source of non-zymasic anaerobic carbon dioxide, i.e. $(I–Z_{CO_2})$, remains unsolved. Amino acids could be a source, but experimental evidence is lacking. Fidler (1951) has reported his Cambridge experiments on apples, concluding that in these fruits most of the CO_2 not attributable to zymasis is formed by the complete anaerobic oxidation of malic acid. The partial reactions which might effect complete oxidation have not been elucidated. Unless plant cells can distinguish between pyruvic acid formed from carbohydrate and from malic acid, this C_3 acid cannot be an intermediate.

Z_{CO_2} is calculated from the accumulation (Z) of ethyl alcohol plus acetaldehyde. $Z_{CO_2} = \frac{44Z}{46}$. The anaerobic consumption of ethyl alcohol or of acetaldehyde (other than its reduction) would lead to calculated values erring on the low side. We have as yet no evidence of such metabolism. Fidler's feeding experiments indicate that in apples and oranges the anaerobic metabolic fate of acetaldehyde is exclusively reduction to ethyl alcohol.

(iii) *Low-oxygen zymasis.* Healthy tissues do not show zymasis in air. Stich (1891) found that below oxygen concentrations of 5 per cent. the respiratory quotients of plant organs rise above unity. Most investigators tacitly assume that rises in R.Q.'s from unity to considerably higher values mean that alcoholic fermentation has set in. Very few make measurements of alcohol accumulation to confirm their assumption.

F. F. Blackman (1928) introduced the term extinction point of anaerobic respiration, for the external oxygen concentration above which respiration is entirely aerobic, and below which it is of a mixed form. From CO_2 measurements he concluded that the E.P. in Bramley's Seedling apple was in the neighbourhood of 5 per cent. Unambiguous evidence of the extinction of anaerobic respiration, when zymasis is a component of this process, can be obtained by demonstrating the suppression of alcohol accumulation. In

1933, Fidler and I reported that alcohol accumulation in apples early in the storage season could be suppressed by *external* oxygen concentrations of 3 per cent. or even less. The E.P.'s for growing apples were about 1 per cent. Ranson found values of 1·8 per cent. for rhododendron leaves and 2·0 and 1·5 per cent. for germinating seeds of marrow and buckwheat respectively.

Evidently a very low *internal* oxygen concentration is adequate to maintain the aerobic consumption of the most degraded product common to both aerobic and anaerobic glycolysis. To explain this fact we require to know the identity of this common product, and also the influence of oxygen concentration on terminal oxidases that control the regeneration of TPN, DPN and other intermediate hydrogen carriers from their reduced forms. Differential retardation within the tricarboxylic acid cycle could lead to the production of the common product faster than it is consumed by aerobic processes. Accumulation of this product would then be easy to understand. However, it is metabolised in a new way, leading to the accumulation of ethyl alcohol. We can describe the partial reactions, but how they are set in train remains a puzzle. Lack of space prevents discussion, which would bear also on suggested explanations of the ' Pasteur effect.'

(iv) *The Pasteur effect.* The essence of the Pasteur effect, when it is displayed, is the power possessed by oxygen of conserving carbohydrate in cells. For illustration, I shall cite some evidence obtained at Newcastle. Ranson compared carbon loss as CO_2 and ethyl alcohol in air, at the E.P., and under anaerobic conditions. He concluded that in all the organs examined except germinating buckwheat, the catabolism of carbohydrate is increased as the oxygen concentration is lowered below the E.P. In other words, raising the oxygen concentration from zero to the E.P., by suppressing zymasis, conserved carbohydrate substrate in these organs. This is the critical test for the occurrence of the Pasteur effect. For some organs (e.g. apple fruits) carbohydrate catabolism increases much more when the oxygen concentration is reduced from the E.P. to zero than when it is increased from the E.P. to the air value. Accordingly in these organs a comparison of catabolism in air and under anaerobic conditions has provided evidence of the Pasteur effect. This is the easiest comparison to make, but it does not give complete information about responses to oxygen. For example, Ranson's results for germinating marrow indicated that carbon losses in air and pure nitrogen were about the same. It was only by measuring carbon loss at or below the E.P. that he obtained evidence of a Pasteur effect. At the E.P. carbon loss was about two-thirds of the loss in nitrogen.

Direct measurements of carbohydrate loss in higher plants have only once been made to compare anaerobic and aerobic catabolism. Fidler (1948) reported his Cambridge experiments, which showed that carbohydrate loss in apples is greater under anaerobic conditions than in air, and thus confirmed conclusions drawn by F. F. Blackman (1928) and Thomas & Fidler (1933) from measurements of catabolic products. We need further data for substrate consumption as oxygen concentration is increased from zero to the E.P., especially for organs such as germinating buckwheat, for which the existence of a Pasteur effect is still in doubt.

Partial Reactions in the Respiratory Consumption of Carbohydrates

(i) *Historical.* The Russian plant physiologist, Palladin, developed the Pflüger-Pfeffer hypothesis that there is a ' genetic relationship between intra-molecular and oxygen-respiration in plants.' He merged the concepts of fermentation and aerobic respiration in a master concept, namely the linkage of glycolytic with oxidative processes, and the coupling of hydrogen and oxygen activation. '*Zuerst wurde freilich diese Brücke nicht für die Tiergewebe, sondern für die Pflanzen geschlagen, und zwar durch Palladin*' (Oppenheimer, 1927). If botanists took the lead, animal biochemists (their name is legion) have amply repaid the debt.

From the experimental fact that the activity of zymase present in a cell-free juice is not affected by oxygen concentration, Palladin and his younger colleague

Kostyschew inferred the occurrence in aerobic carbohydrate catabolism of at least some of the partial reactions of zymase cleavage. They knew that ethyl alcohol does not accumulate in plant cells in air, and believed, justifiably (p. 226), that it is not easily oxidised. Kostyschew pictured oxidation processes, actuated by enzymes and involving oxygen uptake, as incident on a precursor of ethyl alcohol. He found that the respiration of seedlings was stimulated when he fed them with extracts from the residual liquid of alcoholic fermentation. However, his general conclusion was that unambiguous evidence would not be obtained until the intermediate metabolites in alcoholic fermentation had been identified. These are now known, and modern versions of Kostyschew's experiments have been performed with pure preparations of precursors (e.g. acetaldehyde and pyruvic acid) of ethyl alcohol.

During the 1930's it was often stated, without adequate supporting evidence, that acetaldehyde is the substance which links aerobic with anaerobic respiration in the higher plants. The publication of Fidler's report (p. 225) should result in a critical reconsideration of this hypothesis. At present we have no evidence of the existence in higher plants of an acetaldehyde oxidase, i.e. of a dehydrogenase linking with oxygen through a terminal oxidase. Even if acetaldehyde can be oxidised *in vivo*, there is no conclusive proof that it is a normal intermediate in the aerobic respiration of carbohydrate. Evidence exists that pyruvic acid can be oxidatively consumed, but it is possible that the intermediate 2-carbon fragment is acetate, acetyl phosphate, or some other substance, and not acetaldehyde. Once more we see (cf. p. 226) that there is much obscurity about partial reactions involving compounds containing three and two carbon atoms.

Most of the recent work on pyruvate catabolism in plants has related to the theory succinctly described by Professor Sir Rudolph Peters in his Croonian Lecture (1951). Here is a quotation from his lecture: 'it is now believed firmly by biochemists . . . that the 2-carbon fragment formed by oxidative removal of CO_2 from pyruvate is not oxidised directly; but that a condensation with a 4-carbon acid, oxaloacetic acid, takes place to form a 6-carbon tricarboxylic acid; this is then successively degraded by loss of CO_2 through a 5-carbon and 4-carbon stage, until oxaloacetate is again reached; the latter is then ready to pick up another fragment. In one turn of this process, a pyruvate molecule is completely oxidised. This biochemical theory in its present form we owe to Krebs (1943), who with much penetrative generalisation put together earlier schemes by Szent-Györgyi and by Knoop and Martius to make this whole.'

This theory can be regarded as a modern development of Palladin's and Kostyschew's. Phosphorylative glycolysis (see fig. 1), produces pyruvate or a derivative (equation (1)) which is oxidatively consumed (equation (2)) by the tricarboxylic acid cycle, usually so rapidly that it does not accumulate. The respiratory quotient for the complete oxidation of carbohydrate would be unity.

$$C_6H_{12}O_6 + 2H_2O + O_2 = 2C_3H_4O_3 + 4H_2O \quad . \quad (1)$$

$$2C_3H_4O_3 + 6H_2O + 5O_2 = 6CO_2 + 12H_2O \quad . \quad (2)$$

[N.B.—The actual number of water molecules consumed and produced is still problematical. My computation is not offered as a possible solution of the problem. It is based on fig. 1, which ignores phosphorylations and dephosphorylations.]

(ii) *Evidence of stages in intermediate oxidative carbohydrate catabolism.* Some intermediates of zymase cleavage have been detected. For example Bonner (1948) found 0·77 mg. fructose-1 : 6-diphosphate and 0·48 mg. fructose-6-phosphate in 1 gm. fresh weight of oat coleoptile. The accumulation of some intermediates under normal conditions may be too slight for detection. Pyruvate provides an example. James and James (1940) detected an accumulation in cells when a carboxylase inhibitor was added. Morgan (1946) failed to find it in onions until they were minced or ground in a mortar. In my experience no more than traces of acetaldehyde or ethyl alcohol will be found, in the absence of fixatives or of conditions which impair the oxidative power of cells.

Some evidence that the enzymes of

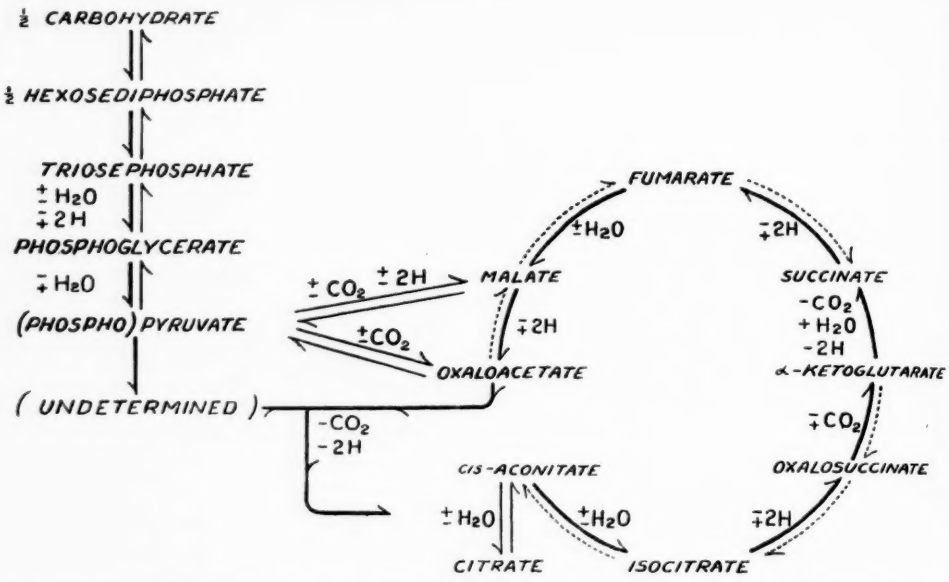


FIG. 1.—The relation between phosphorylative glycolysis (certain steps omitted) and the tricarboxylic acid cycle in the complete catabolic oxidation of carbohydrate. The latest views on entry into the cycle are not pictured or discussed. The development of the theory is left at the stage that a derivative of pyruvate, one of the C_3 fragments of glycolysis, combines with oxaloacetate to yield a tricarboxylic C_6 acid. [In the zymasis component of anaerobic respiration, pyruvate is decarboxylated ($-CO_2$) to yield the C_2 fragment acetaldehyde, from which ethyl alcohol is produced by reduction ($+2H$).] Attention is called to six partial reactions involving dehydrogenations ($-2H$) : one in glycolysis, one in entering the cycle, and four in the cycle. Each dehydrogenation links with side reactions involving O_2 uptake actuated by terminal oxidases : $H_2 + \frac{1}{2}O_2 = H_2O$. There are three decarboxylations leading to CO_2 production ($-CO_2$) : one in entering the cycle, and two in the cycle.

glycolysis are present has come from experiments on cell-free juices or cytoplasmic particles. For example James *et al.* (1944) concluded that barley juice can convert hexose diphosphate into phosphoglyceric acid. Presumably among the components of the juice were aldolase, ATP, triosephosphate dehydrogenase, and DPN (cozymase). Stumpf (1950) reported evidence from experiments performed *in vitro* that aldolase, isomerase, triosephosphate dehydrogenase, enolase, phosphotransferase, and carboxylase, are involved in the enzymic conversion of hexosediphosphate to acetaldehyde by pea seeds. He succeeded in studying each partial reaction separately. He had previously isolated aldolase and purified it ' some 92 times.' After a long delay the difficult task of isolating from higher plants and purifying glycolytic enzymes and coenzymes has begun. However theory is well ahead of convincing experimental proof. For example it is generally accepted that DPN is a key substance, but the only reports of its presence in plants are those of Myrbäck (1937) and Whatley (1951). The evidence for the presence of ATP is stronger. From mung beans Albaum *et al.* (1950) have prepared a specimen of 70 per cent. purity, indistinguishable from rabbit muscle ATP.

All the acids involved in the tricarboxylic acid cycle have been detected in plants but not all together in a single plant tissue at a given time. Under natural conditions in some plants molecules of a certain acid may be short lived. Consumption is so much faster than production that there is little or no accumulation. In other plants this acid may accumulate, reaching relatively high concentrations (e.g. malate in apples, citrate in lemons, isocitrate in blackberry fruits and Bryophyllum leaves).

Feeding experiments have yielded evidence that pyruvate can stimulate CO_2 output and oxygen uptake by slices of

plant tissues. Bonner and Wildman (1946) used spinach leaves previously starved by keeping them in the dark, Bonner (1948) used coleoptiles of seedling oats depleted of carbohydrates by removal of the endosperm 48 hours before the experiment began, Laties (1949) used barley roots, and G. Barron *et al.* (1950) potato slices. Respiration was also stimulated by feeding with certain acids of the tricarboxylic acid cycle. In some of these American experiments, obviously modelled on Krebs and Eggleston's (1940) on minced pigeon breast muscle, feeding was combined with the use of inhibitors. Evidence was obtained that the complete oxidation of pyruvate is inhibited by malonate, and that, with this inhibition, succinate accumulates. If malonate is a specific inhibitor for succinic dehydrogenase (p. 224), the succinate cannot arise from the *reduction* of fumarate. It could arise from the metabolism of pyruvate by the partial reactions of the tricarboxylic acid cycle. None of these reactions would be inhibited by malonate. It has been stated that a crucial test for the occurrence of this cycle is whether or not the *oxidative formation* of succinate from fumarate can be proved. The route would pass through malate, oxaloacetate and so, *via* C_6 acids, round the cycle. Feeding experiments with fumarate on malonate-inhibited plant tissue (i.e. tissue presumed incapable of *reducing* fumarate to succinate) have led to the accumulation of succinate. If the premises of the argument are granted, it must be allowed that the tricarboxylic acid cycle can operate in plants.

There are convincing records of the presence of aconitase, isocitric and malic dehydrogenases, and fumarase, in cell-free plant juices, but none of these enzymes has been isolated and purified. The presence of aconitase demonstrates the possibility of interconversion of the C_6 acids. G. Barron *et al.* (1950) have reported that citrate accumulates when juice extracted from potatoes is added to a mixture containing acetate plus oxaloacetate in the presence of ATP as energy donor, and of malonate as inhibitor of one of the later partial reactions (see p. 224) in the cycle by which C_6 acids are consumed. This experiment provided the first evidence of the enzymic synthesis of C_6 acids by plant tissues.

If plant cells can produce acetate from pyruvate (proof is lacking), the enzymic conversion of a hexose to isocitrate is an acceptable hypothesis. The first step in the regeneration of oxaloacetate is the oxidation of this C_6 acid by isocitric dehydrogenase. This enzyme acts with TPN. The presence of this coenzyme in plants has now been demonstrated (Whatley 1951). The oxalosuccinate produced by this oxidation could undergo β-decarboxylation under the agency of a β-carboxylase. Gollub and Vennesland (1947) detected this enzyme in extracts from parsley roots and other plant sources. Theory requires that α-ketoglutarate, the product of this decarboxylation, should now yield succinate. (Bonner 1948) has obtained evidence that this oxidative decarboxylation occurs in plants, but he did not use cell-free preparations.

It can be maintained that in some plants succinate, acted on by its specific dehydrogenase protein, donates two hydrogen atoms to cytochrome c and is thereby oxidised to fumarate. Hill and Bhagvat (1939), Goddard (1944) and others have reported the presence of succinic dehydrogenase, but the plant enzyme adheres tenaciously to cell particles. From spectroscopic evidence Keilin (1925) concluded that cytochrome c is widely distributed in plants. Goddard (1944) extracted it and cytochrome oxidase from wheat embryos. Without question succinic dehydrogenase and members of the cytochrome system are present in some plants, but it remains to be determined whether or not they invariably occur together. In animal cells they appear to form a single enzyme complex, called succinic oxidase.

Fumarase controls the partial reaction by which water and fumarate combine to yield *l*-malate. The oxidation of malate to oxaloacetate is effected by the malic dehydrogenase, acting with DPN.

The orderly sequence of enzymic partial reactions in carbohydrate catabolism depends upon the efficient regeneration of those hydrogen acceptors, for example DPN, TPN and cyt. c, which are present in low concentrations in cells (cf. p. 5). If a plant contains a terminal oxidase

similar to that in animal cells one atom of oxygen would be absorbed for the conversion cyt. $H_2 \rightarrow$ cyt. Future work must elucidate the nature of the terminal oxidase which effects this conversion in such plant cells as do not contain cytochrome oxidase.

Future work must also determine the hydrogen acceptors which participate in the condensation that leads to the synthesis of C_6 acids, and in the decarboxylative oxidation of α-ketoglutarate. We have some evidence about the aerobic enzymic oxidation to DPN of the $DPNH_2$ formed during the dehydrogenation of triosephosphate to yield phosphoglycerate or of malate to yield oxaloacetate, and about the aerobic enzymic oxidation to TPN of the $TPNH_2$ formed during the dehydrogenation of isocitrate to yield oxalosuccinate. Whatley (1951) has found in leaves an enzymic factor which is concerned with the oxidation of $TPNH_2$.

According to Slater (1950) the components of animal $DPNH_2$ oxidase may be diaphorase I (a flavine enzyme), a factor curiously named BAL (British Anti-Lewisite), cyt. c, cyt. a, and cyt. a_3 (possibly identical with cyt. oxidase), i.e. five distinct partial reactions take place in the regeneration of DPN through the uptake of one oxygen atom. Clearly the continuous oxidation of triosephosphate or of malate is not so simple a process as might appear at first sight. Lockhart (1939) reported the presence in plants of a flavine enzyme which may have been diaphorase I, but knowledge about a plant oxidase involving this enzyme and the cytochromes is lacking.

Higher plants may contain other terminal metallo-oxidases (e.g. polyphenol oxidase, ascorbic oxidase) instead of or as well as cytochrome oxidase. For example the oxygen uptake of potato may be largely owing to the activity of polyphenol oxidase (Boswell and Whiting 1938, Boswell 1945). According to James and Beevers (1950), metallo-proteins do not take part in the very rapid oxygen uptake by aroid spadices. They hint at the possibility that flavoproteins may be acting as terminal oxidases.

Limitations of space have compelled me to omit many references to modern work, but in my view, very little has been really established about the functions of terminal oxidases in higher plants. Possible pathways here, others there, have made a bewildering maze. If, for any organ there is a plan showing the direct way in, through and out, I have not seen it.

(iii) *The oxidative catabolism (or respiration) of carbohydrate as an integral process.* There is justification for the abstraction of some processes, traditionally termed respiratory, from total metabolism. They appear to have some distinguishing properties. Thus a gas exchange quotient, CO_2 produced/O_2 absorbed, approaches unity when carbohydrate is completely consumed. Moreover, there exist balance sheets in which CO_2 output and O_2 uptake are quantitatively accounted for, within a narrow margin, by carbohydrate loss. However, we do not yet possess evidence that, in any visible cell structure (e.g. mitochondria), or in an extractable invisible sub-cellular unit like the zymase complex, there exists a hierarchy of enzymes whose integrated action can effect the complete oxidation of carbohydrate. Green (1951) states that in animal mitochondria there may be a 'physical housing' of the cyclophorase complex of enzymes, i.e. enzymes influencing a complete oxidation of pyruvic acid, fatty acids and several amino acids by the tricarboxylic acid cycle; but evidence is still wanting of the presence in mitochondria of most of the enzymes which actuate the partial reactions of glycolysis. From seedlings of mung bean, Millerd *et al.* (1951) have isolated particles possessing the dimensions of mitochondria and staining brilliantly with Janus Green B. They reported that these cytoplasmic particles could actuate all the reactions of the tricarboxylic acid cycle, and could effect, in the presence of malate, the complete oxidation of pyruvate with the predicted respiratory quotient. However, they presented no evidence that these particles contained the whole assemblage of enzymes which actuate the partial reactions of glycolysis. The complete respiratory oxidation of carbohydrate by plant cells cannot yet be attributed to the powers of a single sub-cellular structure.

There may be over thirty component partial reactions in the complete oxidative catabolism of hexose with an R.Q.

of unity. Obviously, potential in the properties of the whole process are the properties of each separate reaction. In some, hydrogen atoms or their electrons are transferred leading to oxygen uptake. In others carbon dioxide is produced. In these and other reactions there are energy exchanges leading to the liberation of heat or the harnessing in active phosphate bonds of free energy to be used later in driving anabolic events or other cell processes. It has been estimated that well over twenty active phosphate bonds are incorporated into ATP molecules as a result of the complete oxidation of one molecule of hexose.

It is possible that all the properties belonging to single reactions in 'respiration' can be matched by those of partial reactions in other abstracted cell processes, that few physiologists would regard as 'respiratory.' Certainly electron transfer must be a frequent occurrence in the metabolism of aromatic, isoprenoid and many other classes of compounds, and, to some extent, the external manifestations of oxygen uptake and CO_2 output may relate to partial reactions in such forms of metabolism.

It is a first statement only that the respiration of carbohydrate controls a functional process, such as the absorption of potassium ions against a gradient. This first statement when analysed, leads to penetrative questions. In the oxidative catabolism of carbohydrate what are the partial reactions which exercise the control and how do they operate? Attempts are being made to answer these questions thanks to the pioneering experiments in 1933 of Lundegårdh and Burström on anion (or salt) respiration, and to still later work in Sweden, and also to the supporting evidence obtained in other countries, for example in recent years by Robertson in Australia. Steward (1932) reported that the accumulatory mechanism depends upon oxygen supply, and Hoagland and Broyer (1936) found a relation between the sugar content and salt absorbing power of roots. The evidence strongly suggested that respiration provides the energy to move salts against a gradient, and so bring about accumulation. Later work led to the conclusion that 'total aerobic respiration' is not required, but only that part which is inhibited by cyanide, azide, and, according to Weeks and Robertson (1950), in some tissues (e.g. carrot) by carbon monoxide. This last inhibition was removed by light. Accordingly (see p. 224) it was concluded that in certain tissues 'there may be a specific association of the cytochrome system in salt respiration with the accumulation mechanism.' The indications are, therefore, that the performance of osmotic work relates to the properties of partial reactions by which intermediate hydrogen carriers are regenerated. According to Lundegårdh's hypothesis (1940) the essential property is the separation of hydrogen ions and electrons.

This single example must suffice to illustrate the argument that partial reactions may set up the necessary conditions for a physical process to occur inevitably. It may be predicted that, looming increasingly large in the future study of physiology, will be attempts in many directions to relate essential nutritional and other functions to key partial reactions in total metabolism.[1]

References

(1) *Introductory chapters.* Similar ground has been covered in successive editions of my book *Plant Physiology* (Third Edition 1947, J. & A. Churchill Ltd.).

(2) *Crassulacean acid metabolism.* References are given in my articles in the *New Phytologist*, 1949, **48**, p. 390, and in *Symposia of the Society for Experimental Biology*, 1951, number V, 'Carbon Dioxide Fixation and Photosynthesis.'

(3) *Anaerobic Respiration.* The bibliography of this chapter is sufficiently covered in an article on 'Respiration' by J. S. Turner, *Annual Review of Plant Physiology*, 1951, Vol. 2, p. 145.

(4) *Respiratory consumption of carbohydrates.* An extensive bibliography will be found in an article on 'Respiration of Higher Plants,' written by D. R. Goddard and B. J. D. Meeuse in *Annual Review of Plant Physiology*, 1950, Vol. 1, p. 207. Volume 3 of this Review will contain an article by P. W. Stumpf on 'Glycolytic Enzymes in Plants.' In volume 2, p. 1, of this Review, R. N. Robertson deals with the question of the relation between respiration and the absorption of inorganic nutrients.

[1] Delivered Sept. 4, 1952.

EDUCATION OF AFRICANS IN EAST AND CENTRAL AFRICA

Address by

A. L. BINNS, C.B.E., M.C.

PRESIDENT OF SECTION L

I HAVE chosen a somewhat restricted title for this Address because even with this limited heading the subject is big enough, and also because I have recently spent eight months in East and Central Africa, whereas I have paid only shorter visits to North and West Africa. It may be said that eight months is a period all too short to serve as a basis for conclusive statements, and I agree, but on the other hand I had the advantage of seeing the situation in East and Central Africa against a background of long experience in our own countries and other parts of the world. There is no doubt, I think, that there is far too little interchange of administrators and teachers between the British and Colonial Services. Most of those who go from our countries to the Colonies do so in the 20's or early 30's, and most of them remain in the Colonial Service until they retire. It is notoriously difficult for a man or woman to get back into a responsible position in our countries if he or she has been teaching or doing educational administration in the Colonies for any considerable length of time, and few teachers and administrators pass from our countries to service in the Colonies when they are over forty. This appears to me to be a state of affairs which should be altered. Experienced administrators and teachers from our countries would be able to contribute a lot to the Colonies if they went there in sufficient numbers, while many of them would benefit greatly from some years of Colonial experience before settling down to work at home. The Ministry of Education, which has lent its powerful support to the system of interchange of teachers between our countries on the one hand and the British Dominions, America, and European countries on the other hand, seems to me to be due to turn its attention now to the exchange of educational workers between the home country and her Colonies, while Local Education Authorities, generally speaking, ought to show themselves much more willing to give teachers and administrators from the Colonies a chance to show what they can do in really responsible positions in the United Kingdom. I know there are several bright exceptions, but I do not think anyone will seriously dispute the validity of the general statements I have made. Moreover, in our usual British way, we have undertaken in the Colonies vast responsibilities for the welfare of many millions of people who are now rapidly emerging from a primitive way of life, and we cannot discharge these responsibilities adequately unless we send to our African Territories some of our best teachers and administrators, and unless we contrive to find a means of effecting an easy exchange of educational workers between the home country and the African territories. Recruitment of teachers and administrators from our countries for East and Central Africa is likely to prove increasingly difficult because the Colonial Office are rightly determined that quality shall not fall, while the needs of the African Territories for European teachers and administrators are now increasing and may be expected to continue to increase until a sufficient number of Africans have been trained to take over the work, and that will need several more years.

One of the first things that strikes a visitor from Europe is the way in which African education is meeting with many of the problems with which education in our

countries was confronted in the past, and that is only one of the ways in which considerable experience of educational work in our countries could help the Africans. As in our countries, education in East and Central Africa was provided by the Churches in many places and for many years before the State took a hand. Then, as in our countries, the State's contributions were small and tentative in the beginning, and the State has become increasingly involved, sometimes unwillingly and in spite of its efforts to keep out. The Churches having entered the field with the same enthusiasm as in our countries, and having enlisted for that purpose the efforts of many self-sacrificing and earnest men, found their own success in spreading education to be creating a demand for education which they could not entirely meet. Increasingly the State was compelled to provide financial assistance, and with that financial assistance have come problems of control, and sometimes it has seemed that efficiency in secular teaching is at variance with religious freedom. The State has also found it necessary in some places to step in to fill the gaps left by the Churches in covering the country with schools, so that Governments and Local Authorities have provided schools of their own to supplement the Church Schools. Where Governments have been unwilling to enter the field of school provision or to encourage the African Local Authorities to do so, a number of Private Schools have sometimes emerged, again creating problems similar to those in this country, for while some of these Private Schools are deserving of every encouragement, others are inefficient in their teaching and bad in other ways, and parents are undoubtedly failing in some instances to get value for the money they spend in fees, while the wastage of human material is even more serious. As in our countries, some of the African Governments are making preparations to control and inspect Private Schools, and as in our countries in the past, the relations of the Church Schools with the State have presented a number of problems, one of which is very familiar to Educational Administrators here, namely, the maintenance of the fabric of Church School buildings in a satisfactory condition. Other problems of African education read almost like a history of education in the United Kingdom, for example, school fees, selection of children for higher education, transport of school children, school meals, school children's clothing, and that problem which pervades most of the others, the problem of how to finance the ever-increasing cost of the Educational Service. It will not be possible to do much more than mention some of these problems, but there are others which I shall try to examine at some length.

The aim of education is generally agreed to be that of drawing out all that is best in every individual, and we are often told that education should make the individual a good citizen, a good worker, and should enable him to develop all the talents with which Heaven has blessed him. What is sometimes forgotten in our countries as in Africa is that education is never given in a vacuum ; it is always given in relation with the environment of the child, and people must use their education in relation with the environment in which they are placed in adult life. The United Kingdom is primarily an industrial country with an intricate system of administration, not only in its Government but in every part of its commercial life. Administration in industry and commerce is everywhere required at the highest level of efficiency to enable our extremely complicated social and economic machine to work, while a large amount of technical education is needed to meet the demands of industry. The child leaving school usually has the choice of a number of different kinds of work, but in all of them his life is likely to be extremely complicated in relation with a very varied and highly organised environment. In Africa on the other hand, apart from the copper belt of Northern Rhodesia and a few large towns, life is relatively very simple and based on an agricultural economy. The Territories of East and Central Africa depend on their own land to nourish their rapidly increasing population and to furnish the exports necessary to pay for the things that must come from overseas to raise the standard of life. Agriculture is the life-blood of these Territories, but for the most part it is now conducted by very primitive means which are often wasteful not only of the crops but of the

land itself, for soil erosion arising from bad cultivation and over-grazing is a major problem. The number of administrators required by Government and the Local Authorities and by industry and commerce is very small compared with the requirements of the United Kingdom in this respect. The contrast between life in our countries and life in East and Central Africa is, therefore, very marked, and yet the educational aims and methods of the schools of Africa are far too much like those in our countries. For example, the Secondary Schools of Africa usually have an academic curriculum, and they furnish excellent recruits for the African Civil Service, but their products do not return to the land, though it is the land above everything else in Africa which requires intelligent and expert knowledge to enable it to do its work for the people whose very existence depends on it. The problem is to adapt the essentially sound aims of education in the British Isles to the environment of Africa. The nature of that problem is very well understood by the staffs of the Education Departments of these Territories, who are striving hard to solve it, and we may profitably inquire why this problem is so difficult to solve.

In the first place it must be said that the majority of the young teachers who have been recruited from the United Kingdom for service in these parts of Africa have been required to possess University Degrees. That, one would say, is an excellent thing for the University Degree should be and generally is the hall-mark of scholarship, and sound scholarship is needed in Africa as elsewhere. On the other hand, many of the Graduate teachers in our countries are serving in Grammar Schools, and it is, therefore, from teaching service in the Grammar Schools of the United Kingdom that quite a lot of the European recruits for the African Educational Service have come. It must also be remembered that most of the Grammar Schools in the United Kingdom are situated in towns, so that many of the European recruits have come to Africa from an urban environment. In Africa they generally begin work by teaching in Secondary Schools or by assisting in the training of teachers, and there is an obvious danger that their work may begin on urban and academic lines, and that their African pupils and the future African teachers may be educated accordingly, and thus become maladjusted to their rural and agricultural environment. It was most noticeable, however, that many of these young countrymen of ours, coming into a strange land and meeting the problems I have described, have adapted themselves splendidly to the new conditions and have found a way of basing sound scholarship on a curriculum with a rural bias. The work of others is undoubtedly purely academic, but that is not their fault but is due to the demands made on them by the Africans. It is obviously most important, however, that those responsible for recruiting young teachers in our countries for the African Service should always have in mind that the best teacher in Africa will usually be the man or woman who loves the countryside and is capable of transmitting that love to his pupils, or rather of causing them to catch it from him. The minority of Africans who will need to be trained as administrators for the next few years will be all the better if they are trained to administer for a rural people, while the great majority of the Africans who go to school must be educated to live on the land, to make the best of that environment as it is, and to improve it as rapidly as may be. Most of the European teachers and administrators in East and Central Africa fully realise this essential fact, and are doing their best to work in accordance with it. It is no fault of theirs that the problem is still unsolved, and we must look elsewhere for its main solution.

When I was in Rhodesia a prominent African said to me, ' In this country every European is a teacher for the whole of his life, even if he does not realise that fact, and even if he does not mean to be a teacher.' It is undoubtedly true that many of the Africans do consciously or unconsciously model themselves on the Europeans with whom they come in contact, and it must be confessed that very few Europeans in these countries work with their hands on the land or anywhere else. The African who is seeking to live like a European is naturally led as a rule to imitate the external appearances before he grasps the real essentials of the British

way of life, and he finds unfortunately that the shortest cut to the apparent European level is through the possession of a certificate of some kind from school. We often complain in our countries of the tyranny of examinations, but the burden of that tyranny in African education is very much greater than anything we have experienced here in our lifetime, and it is the Africans themselves who have largely been responsible for this state of affairs in a misguided effort to copy Europeans instead of working out their own salvation in their own way. An African often values education for the knowledge it gives him which he can reproduce in an examination paper, and thus obtain a certificate which will get him a clerical job. In other words he values education for the examinable knowledge it gives him rather than for what it enables him to do in life. We cannot blame the Africans, for quite a lot of our own education in the United Kingdom has been directed to the same end, and we must remember that the African's temptations are immeasurably greater than ours. In many parts of East and Central Africa the soil is poor and the climate does almost everything it can to discourage agriculture. Life in an African village is often hard, and indeed primitive to a degree which would appal our own countrymen, but if education is to do its proper work in Africa, as elsewhere, it must enable a great majority of the people to live as well as may be in their own land, and to improve that land as rapidly as possible. In Africa, as elsewhere, education can only enable a small minority of the people to escape from their environment. The problem then is to give the Africans a liberal education with sound scholarship based on the things that exist in the countryside, and which man must use in the countryside, an education based on agriculture and rural industries. Unfortunately, many Africans regard practical work in education as an attempt to fob them off with something cheap and inferior, and we have to try to get them to realise that what we are offering them is what will enable them to live well and their children to live better. We must get them to see that we are not offering a narrow vocational education only suitable for labourers, but an education which will draw out all that is best in them and devote all that is best in them to the betterment of their own country and their own race. There is obviously no short cut to the solution of this problem, but I am confident that it will be solved in time and largely by the persevering and self-sacrificing efforts of administrators and teachers from our countries, trained in sound scholarship and imbued with a love of the African land and the people who dwell on it.

You will gather that I expect it to be necessary to recruit for the next few years an increasing number of young teachers from the United Kingdom for service in Africa. At the same time I want to emphasise that the prime duty of these young teachers will be to enable the Africans to run their own educational system as quickly as possible. In that connection I want now to spend a little time in looking at the typical African teacher of today.

In the first place it must be confessed, I think, that his formal training usually has had to be much inferior to that of his British counterpart, a handicap which is often very severe. In some Territories there is quite a big percentage of completely untrained teachers who have come straight from school into teaching generally after an inadequate education, and have had to make the best they can of their position by learning to teach as they go on, and by improving their own education if they can find means of doing so. We need not be surprised at this state of affairs, for it is only of recent years that untrained teachers have ceased to be a serious problem in the British Isles, and a number of them still exist. The trained African teacher, however, is also usually handicapped by the insufficiency and the relative inefficiency of his training as compared with the teachers in our own countries. The Training Centres for Teachers in East and Central Africa are often small because of difficulties due to religion, language, and communications, and many of the entrants to these centres arrive there with an education which we should consider quite inadequate as a basis for teacher training in our countries. In extreme instances the African student may arrive at his Training Centre able only to speak his own tribal vernacular, and he

may leave the Centre able only to teach in that vernacular. Usually, however, the candidate for the teaching profession learns either Swahili or English as a lingua franca before he comes to the Training Centre, though almost invariably the greater part of his time at the Training Centre is taken up with the furtherance of his own education in English and other subjects, leaving only a comparatively short time for actual teacher training. Here again, we need not be surprised, for even in England a similar situation existed until fairly recently. Efforts are now being made by our countrymen engaged in the Colonial Service of Education to increase the size and improve the efficiency of Training Centres in East and Central Africa, and to transform them into real Training Colleges. The Training Colleges in this part of Africa can never be as large as those in our countries, but an effort is being made to make them large enough to provide for adequate specialist staff, and to enable the students to come into contact with lecturers of varied personality and experience. Too often at present the student comes under the influence of only one or two lecturers, and inevitably copies them far too slavishly in many instances. There is an abundant and very evident need in these new countries for the services of experienced teacher trainers from our own countries who can demonstrate modern Activity Methods of teaching and enable the African teachers to realise that education is a much bigger thing than instruction, and that what counts in life is ability to live well rather than a mere store of memorised information. As it is, African teachers, by reason of their own inadequate education and training, tend to teach too much by formal class methods, and particularly by requiring their pupils to copy notes which they themselves have brought away from their Training Centres, and which the pupils will in due course reproduce in examination papers. I think many of those I am addressing will recognise here also a process which is not unknown in the United Kingdom, even though it is exacerbated in Africa.

But I would like to tell you more about the typical African teacher, now that I have told you of the handicaps under which he starts. There are black sheep amongst them as there are in all vocations and in all lands, but I found most of the African teachers to be thoroughly conscientious and very hard-working. Quite often the Primary Schools have to work double shifts, the same teachers taking one class of children for half a day and a fresh class in the second half. The double teaching time adds up to 7 hours or more in some instances. Quite often these teachers work with very meagre equipment and their salaries are low by our standards, but I was often amazed at what they accomplished, and I could not help thinking how much more they would accomplish with better training and better facilities. Some African teachers are at present reluctant to accept responsibility or to show initiative, and that is quite understandable having regard to the handicaps I have mentioned, but on the other hand we met numerous examples in these Territories of African teachers who were controlling big schools with large staffs in them, and doing that work extremely well. I feel sure that there are great possibilities in this direction, but I am equally sure that these possibilities will only be realised if we can give the African teacher a better start by sending more experienced teacher trainers from this country in the first place, and by raising the educational standards of the students entering the African Training Colleges. Here again the Africans want our help to enable them to stand on their own feet as quickly as possible.

Having said something about the general aim of education in East and Central Africa, and something about the teachers who will be called on to realise that aim, I now devote a little of my time to language considered particularly as the medium of instruction. I have already mentioned incidentally that the training and teaching of Africans suffers to some extent from tribal isolation, and this is particularly true as regards language. Each tribe has its own vernacular, and there are a number of different language roots. Sometimes when the tribe is a big one the vernacular is current over a comparatively wide area, for example, the Luganda language in the Kingdom of Buganda, and the language called Chinyanja which is widely spoken in Nyasaland.

Sometimes neighbouring tribes are able to speak, or at any rate understand each other's vernacular, and sometimes a Territory uses for official purposes only a few of the many vernacular languages spoken within its boundaries, in the hope that these few official vernaculars will in due course become current over the whole Territory. Generally speaking, however, it may be said that the curse of the Tower of Babel has descended with damaging effect on this part of Africa, for often inability to speak anything except his own vernacular is a great handicap to the African who wants to move outside the area of his own tribe, to engage in trade, to become a teacher, or even to learn in school. It seems evident that before real progress can be made every African must be equipped with two languages, his own vernacular and a lingua franca. The vernacular or mother tongue of the African should certainly be encouraged, for in it are the roots of his tribal traditions and loyalties, and it is the language he speaks in the home. The lingua franca should be current over a very wide area. It should be a language with a great literature readily available to the African in cheap books, and it should be a language widely used in commerce and technology. English fulfils all these requirements, and most Africans are eager for an opportunity to learn it. Another lingua franca in use in some Territories is Swahili. This language is derived from a Bantu foundation with many Arabic and Portuguese insertions. It has been widely used on the African coast for several generations, and it has the merit that Africans learn it fairly easily. On the other hand, it is associated in some African minds with slavery; it has an inconsiderable literature, and is unsuited for use in technical education. It is really the vernacular of the coast, and it seems to me that it is so ill-suited to be a lingua franca that it had better be regarded simply as the vernacular of the coast. The importance of a lingua franca in the schools will be readily realised when it is pointed out that most of the vernacular languages cover such a small area that it is not worth while to print school textbooks in them. Consequently a school in which the local vernacular is the medium of instruction often has to work without any reading books at all. Moreover, of the Africans who go to school, the great majority go there for four years or less, and they will clearly forget all they have learnt and in particular they will lose the literacy they have gained, unless their schooling can be followed-up by cheap and abundant reading matter. The number of books and even of newspapers available in most of the vernacular languages is quite small, and very much literacy is now being lost because of the lack of follow-up material. I remember a visit to one school where Swahili was in use as the medium of instruction. That school had just been presented with a library room by a benefactor, and the African headmaster in showing us the room said ' My difficulty is going to be to fill those shelves. There are not enough books printed in Swahili in the whole world to fill half of them.' That seemed to us a striking example of the importance of language in relation to the present situation in the schools of this part of Africa.

I turn now to some of the difficulties of attendance at school which must be overcome by the Europeans and Africans who are working together in the educational system of these Territories. None of these Territories has yet sufficient school places available to allow school attendance to be made compulsory for all children, even with a limited age range, and in the absence of compulsion it is not surprising that attendance at school is often very irregular, particularly when one considers the difficulties which many parents and children have to face. Although immense progress has been made in opening up communications, much of the country is as yet so comparatively undeveloped and the distances involved are so great that children often have to travel a long way to reach a school. Sometimes during periods of food shortage or heavy rain children have to stay at home of necessity, and occasionally the journey to school, or more likely the journey home, is made hazardous by wild beasts. We heard more than one story of a little straggler from the main body of children being picked off by a lion or leopard on his way back home in the evening. As in our countries, juvenile employment often

competes with school attendance, and this is particularly marked among pastoral tribes where children are very useful for herding cattle and goats. Other tribes habitually employ children to keep birds, baboons, and wild pig off their crops. School fees are usually required, and these are sometimes given as an added reason why children fail to complete a school course. School attendance has been compulsory in England for three-quarters of a century and, therefore, we have forgotten some of the difficulties which still beset educators in Africa. During our travels we repeatedly asked administrators, chiefs, missionaries, and teachers, particularly African teachers, how long it was necessary to keep the average child regularly at school under present conditions in order to ensure literacy to a degree that made it likely to be retained in after life. Over 90 per cent. of our witnesses told us that at least four years at school are required for that purpose under present conditions. We asked a further question—' If literacy is not acquired from schooling, is anything acquired by the average African child from his period at school ? ' ' In particular, is there any background or moral effect ? ' About three-quarters of our witnesses replied ' No, if a child does not get literacy he gets nothing,' and some went even further and told us that in their opinion a child who goes to school and emerges illiterate quite often becomes not only disappointed but resentful and maladjusted, feeling that he has in some way been swindled. The consequences revealed by these questions and answers are very serious, for there is in most districts a very serious wastage of children during the first four years of school life. In one Territory there are over 60,000 children in the lowest classes of the schools, and less than 30,000 children in the second classes, and the wastage in other Territories is often almost equally serious. Comparatively few districts are able to get through the first four years of school life half the children who begin school, which means that according to our witnesses at least half the money and effort spent on Primary Education is completely wasted, and of course, the wastage of human material is again even more serious. Those responsible for government and particularly those responsible for educational administration in these Territories are well aware of the existence of this problem, and they are trying very hard in co-operation with numerous Africans to remedy this most unsatisfactory state of affairs, but the problem is much more easily described than solved. Some African Local Authorities have passed an Ordinance requiring every child who begins school to complete the Primary course of four years at least under penalty of fines, and some African Local Authorities are even employing attendance officers to enforce such Ordinances. Such coercive measures, however, often meet with comparatively little success and at present it appears that methods based on persuasion and full information are likely to be more effective. In particular, the African teachers have in some places reduced wastage to a minimum by their personal influence on parents, and that appeared to us to be the most promising line to pursue, though administrators, particularly District Commissioners, African Local Authorities, and African Chiefs can also play a big part, and when their co-operation is fully enlisted they often do so. In its essentials the problem I am now discussing is, of course, very much like the problem of inducing children in our own countries to complete the full course in Grammar Schools, and, as in the United Kingdom, the problem is not likely to be fully solved until parents are sufficiently enlightened to realise all its implications. What makes the problem so extremely serious in these parts of Africa is the emphatic statement of our witnesses that less than four years at school gives a child nothing at all, but I feel sure that here again the problem will be solved in due course by the united efforts of our own countrymen and the Africans working together. Moreover, there are now certain progressive districts in East and Central Africa where the amount of school provision will soon be sufficient to enable the four years Primary course at any rate to be made compulsory. It will, however, be necessary to proceed district by district in this matter because some districts will certainly lag far behind because of the backwardness of their peoples or the natural difficulties of the country, and

we have learnt from our own experience of Day Continuation Schools in 1920 and the following years how difficult it is to enforce attendance in one area while leaving it optional in others. It is particularly urgent, however, to enforce school attendance in the large urban areas of these parts of Africa, for it is in those areas that the African is required to make the most sudden and violent adjustments to fit in with his new environment, and it is in those areas that young people are at present subject to the severest temptations resulting in juvenile delinquency on a large scale ; again a problem with which we are grappling in our own countries. In some parts of the copper belt of Northern Rhodesia numbers of children have become permanently separated from their homes and families and are leading a vagrant and irresponsible life. The Authorities concerned are mustering all their resources to deal with this situation, but compulsory school attendance, as soon as it can be fully provided for and enforced, is likely to prove the most effective solution.

This leads me to the next part of my subject, namely, the general question of morality among young people in these Territories. Contrary to general belief in the United Kingdom, the tribal structure in these parts of Africa included a very definite and on the whole valuable moral code which was sometimes definitely formulated by the Elders of the tribe, but always existent and strictly enforced. By these tribal codes an individual was bound to behave as a loyal member of a community with his rights and duties alike defined for him and deeply ingrained in him both by precept and example. Civilisation, however, with its opening up of communications, its encouragement of migration, its trade and commerce, its cinemas and its wireless, and all its new alignments and orientations, is rapidly breaking up the old tribal system of these parts of Africa, and the old tribal codes of conduct are rapidly receding into the background. There are still some people in our country who tell us that it is a mistake to educate the African ; that he is happiest in his native ignorance and should be allowed to remain in that condition. Apart altogether from considerations of religion and humanity, the advice thus given is based on a complete misunderstanding of the position. It is out of the question to leave the African alone, for civilisation has already altered his situation out of all resemblance to that of even twenty-five years ago and the process is continuous. If the formative influences of education are not brought to bear on the African there is a real danger that the disruptive influences which the white man's ways always bring into African life will bring about progressive demoralisation. In fact, it is no exaggeration to say that the situation really resolves itself into a race between education on the one hand and disruption on the other. Personally, I believe that no moral code that is worthy of the name can be built except on a basis of deep and sincere religious belief, and therefore, I most earnestly hope for a widespread system of education based on religion. Even those who would disagree with me here will have to admit, I think, that the peoples of these Territories must be educated in the best sense of that word, and as quickly as possible if disaster is to be avoided. I would like to see Government putting all its authority and all possible effort behind the Churches to ensure that a really sound religious education is made the foundation of a widespread and efficient system of schooling whether the schools are provided by Governments, Local Authorities, or Missions.

I have already referred incidentally to the problem of finance, which pervades all others in these countries as in our own. In the United Kingdom the Central Government raises by taxes, I am told, over £80 per head of the population. In Northern Rhodesia the revenue of the Central Government approaches £10 per head of the population, but in Nyasaland, for example, the corresponding figure is only just over £1 per head, and the other Territories all raise less than £10 per head of the population, and most of them considerably less, by central taxation. In some Territories the African Local Authorities also raise a certain amount of money just as the Local Authorities do in England, though not nearly so much. It will be realised that these African Territories are very poor at present by our standards, and that an increase of productivity and a development of their natural resources is

essential if their social services are to be greatly expanded. It will also have been realised, however, that any considerable increase of agricultural productivity must depend on the right education, and thus we have the makings of a vicious circle.

I will not pose as a financier, but it seems quite clear to me, and I hope it will be generally accepted, that these young countries can scarcely be expected to finance the immense developments which they need from their present limited resources. Grants made under the Colonial Development and Welfare Acts have already helped a lot, but they have not proved sufficient to break effectively into the vicious circle which I have indicated above. For various reasons it is difficult for these undeveloped countries to raise substantial loans, though it seems to me that they will be fully justified if they can do so, in financing much of their expenditure on development from loan, since the advantages accruing from present-day expansion will benefit future generations in these Territories when the population will be larger and the productivity much greater, so that repayment of any loans now raised should fall upon a more numerous and richer people. Perhaps we could best help the Africans for the time being by a system of guaranteed loans to enable them greatly to extend their educational provision, including technical and agricultural education, soil conservation, forestry, and animal husbandry. Apart altogether from technical education widespread primary education would undoubtedly help to increase productivity immediately, for one Agricultural Officer said to me, ' If only the Africans in my district could read a simple notice I could reach so many more of them than I can influence now.' That man had gained the confidence of the Africans in his area, but his means of communicating with them were limited to an extent which is difficult for us to imagine by lack of literacy and by difficulties of vernacular.

Most of this Address has been devoted to a description of the vast amount of work we still have to do to enable the Africans to stand on their own feet and become effective partners with us in these new countries, but I want to close on another note. In most of these lands the educational work of the Missionaries has only been sufficiently extensive to make any real impression for some thirty or forty years, while the connection of Government with education is even more recent and in some Territories has only made itself massively felt since the last war. Our countrymen in these parts of Africa are few in number, and their resources are very small by our standards. It is true to say, I think, that never in the history of education has so much been done by so few in such a short time, and with such limited resources. I would like to assure those of you who have not been to Africa that we have every reason to be proud of what the Missionaries and the Colonial Service have accomplished. My colleagues and I have tried to make some constructive suggestions for future development in the Report which we have made to the Governments concerned, but our overwhelming feeling is one of respectful admiration for the ability, hard work, and self-sacrifice of our countrymen in these lands, and admiration also for the great mass of Africans amongst whom they have worked, and who in many instances are making a magnificent response. It would be idle to deny that there is a certain amount of political agitation among the Africans. Taking into account the large amount of freedom which we British allow, it would be amazing if no grumbles were heard, and it is undoubtedly right that anybody who wants to grumble should be allowed to do so. On the other hand, I was immensely impressed by the smallness of the numbers of those who are making a living by deliberately creating ill-will. Where such persons are successful their success is almost always based on ignorance. Goodwill already exists on all sides and in plenty, and everywhere ignorance is the chief foe and the chief deterrent to progress. Widespread education of the right kind will prove the salvation of these Territories and the solution of many of their problems. It will enable their peoples to raise the standard of their living, to adjust themselves to new conditions, increasingly to export their produce to the rest of the world, increasingly to govern themselves and progressively to raise their moral stature.[1]

[1] Delivered Sept. 5, 1952.

BRITAIN'S FOOD SUPPLIES

Address by

DR. N. C. WRIGHT

PRESIDENT OF SECTION M

AT the Inaugural Meeting of this Section in 1912 the late Sir Thomas Middleton, who was the first of my predecessors in the Presidential Chair but who was at the time an Assistant Secretary in the Board of Agriculture, prefaced his Address with the following words :—

'A President of a Section is expected to speak about his own work ; but mine belongs to that category which many people seem anxious to learn about before it has been made public, and in which nobody is particularly interested once, in buff or blue or white garb, it has been issued by His Majesty's Stationery Office.'

Sir Thomas therefore prudently avoided dealing with current agricultural affairs and devoted his Address to a stimulating exposition of the work of the early associations for promoting agriculture and, as he put it, for 'improving the improver.' Although I find myself in the same dilemma I have eschewed his caution and have set myself the task (albeit still, I fear, within the ambit of official publications and pronouncements) of reviewing the past trends and present position of the nation's food supplies. I have done this for three reasons. *First*, because it seems appropriate at a meeting held in Belfast to review some of the outstanding achievements of British agriculture, to which Northern Ireland has made such notable contributions both materially and administratively ; *second*, because in implementing these achievements, and indeed in the parallel task of planning the nation's overseas procurement policy, advancements in scientific knowledge (which form the cardinal feature of the Association's interests) have played a vital part ; and *third*, because the importance of our food supplies, and in particular the importance of increasing still further the home output of agricultural products, has never been so great or so universally recognised, and the lessons to be learnt from recent past experience never, therefore, so relevant.

It is, I feel, particularly appropriate that this Section should be used as a forum for a survey of the nation's food supplies, for the Association has always taken a leading part in relating nutritional needs to national life. To go no further back than 1880, the Association broke new ground by appointing a Committee whose terms of reference were 'to inquire into . . . the present appropriation of wages and other sources of income and to consider how far these are consonant with the economic progress of the people of the United Kingdom.' This inquiry resulted in a pioneer survey of national expenditure, in which expenditure on food was given pride of place. The Report of the Committee has, indeed, since been commonly used as an appropriate starting-point for discussions on the nation's food requirements and supplies.

The Belfast meetings of the Association have themselves been especially notable for the interest displayed in food and feeding. At the meeting held here one hundred years ago, J. B. Lawes and J. H. Gilbert contributed an outstanding paper on the composition of foods in which they showed that prime importance should be attached to the energy, or calorie value, of the food intake. While their experiments were mainly conducted on sheep and pigs, they referred to confirmatory trials carried out on 'hunting horses, foxhounds, pugilists and runners.' It is significant to note that their conclusions, secured by observations on such a galaxy of sporting talent are, within the limits of

nutritional knowledge then available, largely in harmony with the most modern concepts of human food requirements.

Fifty years later Professor H. E. Armstrong, in his Presidential Address to the Section of Educational Science at the 1902 Belfast meeting, confirmed these views in his inimitable way.

> 'I may take food', he stated, 'as another subject of which we fail to discern the importance and which is outside the schoolmaster's ken—although teachers have stomachs as well as other men, and boys in particular are believed to take some interest in that organ. *The subject of food is but a variant on that of energy, for the food we take is mainly of value as the source of energy we expend. . . .*'

to which he incidentally added, in typical Armstrong fashion and as a plea for the inclusion of cookery in the school curriculum,

> 'From food to the preparation of food is an easy step; in point of fact the knowledge of how to prepare food is of far more importance than any knowledge of what food is or does, as on it depends much of the happiness and health of mankind.'

Most of us would not, I think, be prepared to go the whole way with Armstrong; knowledge acquired since his day has indicated the need for choosing the right foods as well as for serving them tastefully. But his two main contentions embody principles to which most students of nutrition would subscribe—that the energy content of the diet is its most important single attribute, and that the palatability and attractiveness of foods are prerequisites to their adequate consumption.

At the same meeting Dr. Edwin Cannan, in his Presidential Address to the Section of Economic Science, stressed the importance of another problem touching on the nation's food supplies. He referred to the rapidly expanding urban areas where, he said, 'the population are so packed that they cannot grow anything for their own consumption; yet every morning their food arrives with unfailing regularity, so that all but an infinitesimal fraction would be extremely surprised if they did not find their breakfast to hand. It is clear,' added Dr. Cannan, 'that there is in existence some machinery of procurement which, in spite of occasional failures here and there, does its work on the whole with extraordinary success.' Dr. Cannan, speaking in a period of free world trade, attributed this success to economic incentives; it is one of the problems of our own era to achieve equal success in the face of even greater urban concentrations of the population, in a period of increasing food shortages, and in conditions of acute currency disequilibrium.

It was, again, at the 1902 Belfast meeting that the Council of the Association appointed a Committee under the Chairmanship of Professor C. S. Sherrington (who, as Sir Charles Sherrington, was subsequently to occupy the Presidential Chair) to report on 'the conditions of health essential to the carrying on of the work of instruction in schools.' It is significant that among the eight criteria selected by the Committee as appropriate for consideration, bodily nourishment was placed first on the list.

I have deviated somewhat from the main theme of my Address in order to show the degree of interest in nutritional matters taken by members of this Association up to the date of the last Belfast meeting. A milestone in the Association's history, so far as food production interests are concerned, was taken ten years later, when the present Agricultural Section was inaugurated. It might have been anticipated that the establishment of the new Section would have been reflected in a greatly quickened interest in the relation of home agricultural production to national nutrition. This expectation has not been fulfilled. It is true that at the inaugural meeting Sir Henry Rew contributed a paper on the sources of the nation's food supply, and that he followed this three years later by a Presidential Address under the title of 'Farming and Food Supplies in Time of War.' Thereafter while the Presidential Addresses to the Section have covered a wide range of subjects, each of which has been germane to some individual aspect of the country's agricultural production, few if any (so far as I can trace) have dealt with the wider aspect of the relation of the country's total food supplies to its total needs. I should hasten to add that this shortcoming (if so it may be termed) does not apply to the consideration of *world* food needs. Ever

since the Presidential Address given to the Association in 1898 by Sir William Crookes (an Address which Sir John Russell characterised as ' one of the most momentous the Association has ever heard ') there has been continuous interest in the world picture. In our own Section Sir Daniel Hall dealt in his Presidential Addresses in 1914 and 1926 with the subjects of land usage and population pressure, in 1920 Sir Frederick Keeble discussed the relationship between population and cropping intensity, in 1929 Sir Robert Greig reviewed Commonwealth agricultural production, and in 1948 Dr. G. Scott Robertson chose for the title of his Address that of ' Agriculture and the World Food Problem.' These surveys culminated in the masterly review of 'World Population and World Food Supplies,' which formed the subject of Sir John Russell's Presidential Address to the Association four years ago. In the light of this interest in the world picture, it is surprising that we have devoted so little attention to our own national needs and to the ways in which these have been met.

I do not propose in the present Address to attempt any detailed historical picture of Britain's food supplies, but rather to deal with recent trends and problems. Nevertheless it will be desirable at least to summarise briefly the general background to the immediate pre-war position.

During the nineteenth century our substantial industrial development had made it possible for us to draw to a greatly increased extent on overseas sources for the supplies of food needed for our rapidly expanding population. Between 1840 and 1880, for instance, our imports of sugar had increased fourfold, our imports of wheat fivefold, and our imports of butter, cheese and eggs nearly sevenfold. Our imports of bacon had risen from nil to nearly 16 lb. per head per year, and only the lack of development of refrigerated transport had prevented our carcase meat supplies from showing a similar striking increase. In the fifty years between 1880 and the 1930's this trend had continued, with only a slight reverse during 1914–1918, until in the immediate pre-war years we depended on imports to the extent of some 70 per cent. of the calorie value of our total food supplies. The extent of this dependence varied, of course, with different commodities. It was highest for cereals, which could be produced cheaply on the *extensively* farmed areas of the New World, and for such crops as sugar, oilseeds and nuts, and citrus fruits, for the production of which a tropical climate is favourable. Thus we imported 88 per cent. of our wheat and flour, 84 per cent. of our sugar and of our fats and oils, 75 per cent. of our fruit and nearly 70 per cent. of our animal feeds. Our dependence on imports was less marked, though still high, for animal products, where our mixed farming methods and favourable grazing climate placed us on a more competitive basis with overseas producers. Thus our imports of animal products varied from 75 per cent. for cheese and 66 per cent. for bacon to 50 per cent. for carcase meat and 35 per cent. for eggs. Our dependence on imports was lowest for those foods which are highly perishable, such as milk, potatoes and vegetables, where we relied almost wholly on home production.

It was, of course, largely due to the exceptionally high import figures which I have just quoted that we were able to enjoy such a remarkably varied diet in the immediate pre-war years in spite of the theoretically high population pressure on the land. Thus, in spite of a population increase of practically 50 per cent. between 1880 and the 1930's, our consumption of meat *per head* per year had increased from 90 to about 130 lb., of fats from 12 lb. to nearly 50, of eggs from 11 lb. to nearly 30, and of sugar from about 70 lb. to nearly 110. At the same time we had reduced our consumption of both flour and potatoes from just under 300 lb. to just under 200 lb. These changes had, however, been paralleled by a regression in our home output of food. Total arable acreage had fallen by a third, wheat by a half, barley by two-thirds and field beans by three-quarters. And although there was some compensating expansion in the cattle population (which was itself partly offset by a reduction in the number of sheep), this amounted only to about 33 per cent. over the 1880 level—compared with an increase of more than 100 per cent. in the total consumption of meat.

These figures may be taken as some indication of the degree to which, in the immediate pre-war years, we had allowed ourselves to become dependent on food supplies from overseas, and of the extent to which we had neglected the food producing potentialities of our own land.

The outbreak of war in 1939 brought us up against these hard realities. It became clear that shipping difficulties would constitute the most immediate and the most serious menace to our food position. The extent of the reduction which took place in food shipments during the war is not perhaps always realised. Before the war such shipments (including both human foods and animal feeds) amounted to an annual total of over 22 million tons. By 1940 this had fallen to 19 million tons, by 1941 to 15 million tons, and for the period 1942 to 1944 to about 11 million tons—barely half the pre-war figure. Against this background of drastically reduced overseas shipments we had, if we were to sustain the war effort, to provide food supplies adequate to ensure the maintenance of the health, morale and productive capacity of a population of very nearly 50 millions.

From the nutritional aspect I think I might summarise the three main objectives in the planning of our war-time food supplies in the following terms :—*First*, to ensure that the population were provided with adequate energy-producing foods—that is to say, to ensure an adequate calorie level in the diet. This has been shown repeatedly (ever since the outstanding work of Lawes and Gilbert to which I have already made reference) to be the most important single factor in maintaining a nation's morale and working efficiency. *Second*, to furnish adequate supplies of ' protective ' foods, not merely as a means of safeguarding general health, but of meeting the special nutritional needs of certain particularly vulnerable groups. And *third*, so far as possible to provide, within the resources available, a diet which was reasonably varied and palatable, since the amount of food consumed, and therefore its effectiveness in ensuring the maintenance of the nation's morale and working efficiency depends (again as Armstrong had so clearly indicated) ultimately on its acceptability by the population. I propose to deal with these three objectives *seriatim*.

The most important step towards attaining the *first nutritional objective*, that of ensuring an adequate calorie level in the diet, was to re-orientate our home agricultural policy in such a way as to increase very substantially the acreage under arable crops. The primary purpose of this re-orientation was to produce the maximum practicable output of calories for direct human consumption. Consequently the area under wheat, which in 1880 had stood at 3¾ million acres but which by 1934–1938 had fallen by nearly 2 million acres, was increased until, by 1943, it had again all but reached the 1880 figure. The area under potatoes was similarly increased from a pre-war level of about ¾ million acres to nearly double this figure, while sugar-beet (the production of which had been substantially increased before the war, in part at least as a defence measure) was raised by a third up to a figure which allowed our existing beet factories to run to full capacity. Simultaneously the numbers of pigs and poultry which, as large cereal consumers, are not only direct competitors with man for food but which waste, in the process of conversion, nearly six-sevenths of the calories consumed, were substantially reduced. It is true that to offset the very drastic cuts in imported feeds the areas under oats, barley and beans were, like wheat and potatoes, markedly increased ; the tendency was, however, to divert these feeds to ruminants, and particularly to dairy stock, which were capable not only of securing the bulk of their nutrients from the available grass and coarser fodders, but of furnishing as an end-product a food of specially high ' protective ' value.

An increase in the consumption of liquid milk constituted, indeed, one of the chief means of maintaining the supply of ' protective ' foods which I have noted as the *second nutritional objective* of our war-time food policy. The war-time maintenance of milk production was itself no mean feat in view of the curtailed grassland acreage, the reduced availability of the most suitable cattle feeds, and the shortage of skilled farm labour. By increased sales off farms and the planned diversion of milk from manufacture to the

liquid market the quantity available for direct consumption was, however, raised from the pre-war level of 820 million gallons to over 1,200 million gallons—an increase of practically one-half. Simultaneously, and particularly to ensure that this increase found its way to the more vulnerable groups, the sale of milk was highly subsidised, both generally and by special incentives under the National milk and the milk-in-schools schemes. In this respect the policy in the Second World War differed fundamentally from that in the First, when the essential nature of ' protective ' foods was still unrecognised and milk consumption was allowed to fall by 25 per cent. But other steps were also necessary to maintain the supplies of ' protective ' foods. The drastic cuts in citrus imports had markedly reduced one of our more important pre-war sources of vitamin C, while supplies of vitamin A had been curtailed by similar drastic cuts in butter. To offset these substantial losses steps were taken to increase the acreages of compensating vegetable crops. The areas under greens were doubled, under carrots nearly trebled, and under tomatoes quadrupled. Perhaps I should add that the acreage under hops, which also provide a protective (or at least a palatable) product was also slightly raised ! These acreage changes were, moreover, supplemented by other steps to increase the ' protective ' value of the nation's staple foods, notably by raising the extraction rate of wheat (a step enforced by supply difficulties, but which resulted in a more nutritious loaf) and its reinforcement with added calcium, and the addition of vitamins to margarine which, as a result of the butter shortage, had become an increasingly important source of fat.

The third nutritional objective, to furnish the available nutrients in a form which would be acceptable to the population, assumed increasing importance as the diet became increasingly monotonous and ' starchy ' in character. Before the war some 34 per cent. of our total calories were derived from grain and potatoes, by mid-war this figure had risen to 43. On the other hand, before the war, meat, fish, eggs, and fats furnished 37 per cent. of the total calories ; by mid-war this figure had been reduced to 30. Such changes may not appear large when compared with the differences between, for instance, typical Western and Eastern diets. Nevertheless, in terms of the pattern of meals which the Western housewife is accustomed to serve, they were sufficiently significant to necessitate palliative measures. These measures took a threefold form. *First*, steps were taken so to adjust the country's import programme that this would include a higher proportion of the more palatable and concentrated foods. Thus, increasing quantities of canned meats and fish, pork and bacon, cheese and (at a great saving in shipping space) dried egg and dried milk were imported, chiefly from the United States and Canada and largely as Lend-Lease shipments. The special value of these shipments is illustrated in the fact that whereas, by 1944, they furnished 10 per cent. of the total calories of our diet, they provided no less than 18 per cent. of the animal protein and 20 per cent. of the fat. *Second*, particular attention was given to the provision of attractive ' spreads ' by the selective allocations of scarce ingredients, notably by specially high allocations of sugar for jam and other preserves. At the same time the food manufacturing industries took all practicable steps to ' extend ' their allocations in such a way as to produce the maximum output of cake confectionery and similar variants of the diet. *Third*, encouragement was given to the increased home production of the more highly flavoured crops. Of these, onions were the most notable example, the area under cultivation being increased from a pre-war figure of some 2,000 acres to over 16,000 acres in 1944.

I am only too well aware that, in attempting to list some of the more important measures adopted to implement the three main nutritional objectives of our food production and procurement programmes, I have over-simplified what is in fact an extremely complex picture, and that I have omitted to make reference either to the broad political and international considerations which finally governed the nation's food policies or to the administrative machinery which was essential to their implementation. My object has, however, been rather to

indicate, sometimes in very broad outline and sometimes by reference to seemingly minor detail, that nutritional considerations *did* play a vital part in ensuring the adequacy of the nation's war-time diet. The extent to which these objectives were achieved may be judged by reference to the figures for the nutrient content of the diet between 1939 and 1944. Pre-war food supplies had furnished a national average of 3,000 calories per head per day. The initial result of the impact of war was to reduce this figure to levels between 2,800 and 2,820—levels which are below what we now recognise to be desirable for the maintenance of a nation's full working efficiency. Nevertheless, as the measures taken to raise home production and to re-orientate our overseas procurement policy took effect, the calorie level rose until, by 1944, it actually exceeded the pre-war figure. Parallel with this improvement, indeed in some respects preceding it, there were all-round improvements in the levels of the individual nutrients of the diet above those of the pre-war period. Thus the total protein content of the diet rose from 80 to 87 g. per head per day, the calcium (of which, on the basis of accepted requirements, there had been an overall deficiency in the immediate pre-war years) increased from under 700 mg. to over 1,000 mg. per head per day, and there were correspondingly favourable increases in the levels of iron, of vitamin C, and of the vitamins of the B complex, the pre-war intakes of which had in most instances been marginal. From the nutritional aspect these figures indicated therefore not merely the maintenance of, but substantial improvements over, pre-war levels.

The measure of the contribution of British agriculture to these improvements is reflected in the fact that they were achieved in spite of drastic cuts in our import programme. Thus we had to face a reduction in wheat imports of 2 million tons, in edible grains and pulses of a million tons, in sugar of well over a million tons, in fruits and vegetables of some 2 million tons and in animal feeds of no less than 5 million tons. To compensate for these losses we produced at home a million and a half tons more wheat, a million tons more beet, a million tons more oats and barley, over a million tons more fruit and vegetables and practically 5 million tons more potatoes. Nevertheless, in spite of these vast increases in output, it is pertinent to note that we only raised our home contribution to the nation's total food supply (as measured in terms of calories consumed by the population) by one-third, i.e. from 30 to 40 per cent., which gives an indication of the stupendous effort that would be entailed in achieving a really substantial measure of self-sufficiency in food.

Moreover, these increases were only achieved at great sacrifices to our livestock industry. The pig population was more than halved, poultry were reduced by a third and intensive poultry keeping virtually eliminated, sheep were reduced by a fifth and were (as a result of the subsequent severe winter of 1946) to reach the lowest recorded figure for more than a century and a half. Only milking cattle showed any expansion, but even so at the cost of a decrease in the supplies of home-killed beef.

I have already stated that throughout the war the outstanding problems of food procurement were due to shipping shortages. The immediate post-war period brought the nation face to face with new but no less urgent problems, though of a different nature. The first effect of the cessation of war was to create greatly increased demands on the world's available food surpluses to meet the urgent needs of the liberated, but agriculturally disorganised, areas of Europe and the Far East. Superimposed on these sudden and substantial post-war demands on supplies hitherto freely available to us, this country had in 1945 to face the further serious problem of the end of Lend-Lease aid, on which we had relied so largely for our war-time food.

To have met these dual problems by reducing either the quantity or the quality of the national diet would, to say the least, have been to involve ourselves in a constantly descending standard of living, since a high level of industrial productivity can, particularly under peace-time conditions, only be anticipated if there is both an adequate and an acceptable food intake. Since we had been forced during the war to rely largely on imports for our

supplies of animal products, and since these represented a most expensive drain on our currency, and particularly our dollar resources, the main objective of the 1947 Agricultural Expansion Programme was to re-expand our livestock industry, but to do so on the basis of substantially increased dependence on home-produced animal feeds and without any appreciable reduction in our total self-sufficiency. Only by this means could we hope to maintain the level of our war-time contribution to the country's food supplies, while simultaneously providing the population with the type of diet to which it was accustomed and which was considered essential to the maintenance of national well-being.

This task, and the extent to which it has since been achieved, can perhaps be stated most concisely by considering side by side the contributions made by British agriculture to the energy value of the diet (in terms of calorie content) and to the net output from the land (in terms of monetary value). Between 1939 and 1943–44 the contribution of home production to the calorie content of the diet had, as I have already indicated, increased by about one-third. On the other hand, during the same period net output, measured at constant prices, had increased by only about a fifth to a quarter above the pre-war level. By 1950 although the calorie contribution had been maintained at the same raised figure, the net output had increased to over two-fifths above the pre-war level. Thus while the re-orientation of British agriculture during the war involved an expansion in home-produced calories at the relative expense of the net value of our agricultural output, in the post-war years a substantial increase in net value was achieved without any proportionate loss in calories, i.e. in overall self-sufficiency.

This picture might be easier to appreciate by reference to individual commodities. By 1950 our production of wheat was still a million tons higher than pre-war, our potato output was nearly twice as high, and our sugar-beet was over half as high again. Our beef production was slightly higher than in the pre-war years and our egg production substantially higher. Although the output of pigmeat still fell short of the pre-war figure, it had doubled since the worst war years, and the recovery of mutton had only been delayed by the disastrous winter of 1946. Simultaneously the production of milk had shown consistent annual expansion and had, by 1950, risen to 30 per cent. above its pre-war level. These increases in the output of livestock products were secured on a grassland area which was nearly 5 million acres less than in the pre-war years. Moreover, although the imports of feeding-stuffs had increased since the mid-war period, they still represented only about one-third of those pre-war, the difference being largely bridged by proportionate increases in the home production of coarse grains and the conservation of grass herbage for winter use.

It would be misleading to ascribe to these increases in home production the whole of the improvement which took place in the diet between the mid-war years and 1950. Nevertheless, although it is true that the imports of a number of commodities had been substantially increased, few had reached and still fewer had exceeded the pre-war level. It should, moreover, be pointed out that between 1939 and 1950 there was an increase in our population of some 2½ million for whom additional food had to be secured.

It is therefore in large measure a tribute to our home agriculture that within five years of the end of the war the diet not only showed a marked improvement in comparison with the *mid-war* years but that in all but two commodities, meat and sugar, it stood comparison with the *pre-war* levels, and in respect of two commodities, milk and eggs, surpassed them. The diet showed, too, a welcome return to a less 'starchy' pattern and to greater variety and acceptability. Simultaneously it had preserved the higher nutrient content which was such an outstanding feature of the war-time years.

I have already indicated that the largest single factor responsible for the war-time increase in our food supplies was the change in land usage to a pattern which yielded maximum nutrients for direct human consumption. Equally the post-war expansion derived its success mainly from the superimposition on this pattern of an expanded livestock population

deriving its nutrients largely from home-grown feeds. These re-orientations in agricultural pattern were, however, accompanied by technical improvements in agricultural practice which were designed to increase both yield and productive efficiency. Without attempting to deal with each in detail it would, I think, be fitting at a meeting of this Association, which is so closely concerned with the applications of new scientific knowledge, to refer at least briefly to some of them, preferably in quantitative terms.

First, our agricultural output was aided by very substantial increases in fertiliser usage. Nitrogen usage, which in 1939 stood at 60,000 tons, increased to 180,000 tons by 1944 and to 220,000 tons by 1950. Over the same periods the tonnages of phosphates, potash and lime showed parallel increases, being roughly doubled by 1944 and trebled by 1950. Moreover, not only were these fertilisers applied on an increasingly extensive scale, but during the war years rationing and restrictions on their use tended to ensure their application to the best advantage.

Second, the expansion of the tillage acreage was facilitated by a marked increase in farm mechanisation. The number of tractors rose from the pre-war figure of 55,000 to nearly 300,000 in 1950. Combine harvesters showed an even more striking increase, from a mere 150 before the war to 13,500. By 1950 over 2 million tons of grain were in fact combine harvested. Mechanised methods were similarly developed in other agricultural fields, the most recent and most notable being in the harvesting of sugar-beet. Such developments not only facilitated cultivation and harvesting but were designed (though through lack of experience with incomplete success) to maximise output *per man* as well as *per acre*.

Third, the expansion programme was increasingly aided by modern crop protection methods. The dressing of seed with fungicides had long been a recognised practice ; by 1950 it was estimated that no less than 80 per cent. of the country's wheat and 50 per cent. of other cereals were so treated. This practice is still expanding and to it has now been added the new technique of combined seed dressing which is designed to confer protection against wire-worm as well as fungus growth. Still more striking has been the large-scale development of herbicides which, as spray or dust, are now estimated to be employed for weed control on more than a quarter of our cereal acreage. In the field of animal production similar protective or curative measures were employed. Strain 19 was used extensively to reduce the ravages of abortion, while penicillin treatment undoubtedly curtailed the losses from mastitis. And although it is difficult to estimate its direct effect on meat or milk production, the number of attested herds rose from a pre-war figure of 1,500 to 60,000, so that by 1950 the proportion of attested animals had reached over a quarter of the cattle population.

Fourth, marked improvements were effected in pasture husbandry and grassland management particularly as a result of the very considerable expansion in ley farming, and great strides were made in the development of means for conserving surplus grassland herbage. In particular the output of silage was increased from less than half a million tons in 1946 to nearly 2 million tons in 1950, while within the same period the production of dried grass rose from below 50,000 tons to nearly 200,000.

Fifth, consistent efforts were made to raise the productivity of both crops and stock. Most notable were the steps taken to raise the yields of milking cows. The number of milk recorded animals doubled, and by 1950 covered a fifth of the total population. Still more striking, particularly in view of its comparative novelty, the number of approved bulls at artificial insemination centres increased tenfold, and by the end of the period under review the number of inseminations from such centres totalled more than half a million a year.

The effects of these and other modern applications of science on the country's agricultural output is not easy to assess. Crop yields are recognised to have increased by 10 per cent. since pre-war years, and grassland yields by at least the same proportion. Milk yields are estimated to have risen by two to three times this figure. But whatever the overall contribution, the stimulus to new developments and new techniques engendered

by the war, and throughout the immediate post-war period, should bear an increasing harvest for many years to come.

May I turn now to a brief discussion of the country's future food requirements and to the extent to which British agriculture may be called upon to meet them? This latter will depend not only on the total level of demand for food, but also on the volume of potential overseas supplies.

Three factors will govern the total level of demand for food. *First*, the growth in population. I have already noted that in the decade between 1939 and 1950 food had to be provided for an additional 2½ million mouths. The rate of population growth is slowing up. Nevertheless, over the next twenty years an increase of a further 5 per cent. is still forecast, adding another full 2 millions to the claims on our supplies. *Second*, the extent and pattern of demand. It seems unlikely that, on a *per caput* basis, the former will fall below the pre-war level. As regards the latter, it is noticeable that the trend in the national demand for food is already showing a return towards the pre-war pattern, and although it may be possible to vary slightly the choice within individual classes of food, it is doubtful if the population would—unless in dire necessity—contemplate with complacency a permanent change in general pattern of the type necessitated in the mid-war years.

There is, indeed, evidence that the demand for the more nutritious, palatable and attractive foods may in the future show a tendency to rise. Before the war dietary surveys indicated that there were marked differences in food consumption levels between different income groups within the population. By 1950 these differences had either been eliminated or markedly narrowed. Thus, while pre-war milk consumption figures showed variations from 1 pint per week in the lowest income group to nearly 6 pints in the highest, by 1950 these extremes had been narrowed to 4½ and 6¼ pints respectively. Again, for eggs the pre-war surveys showed extremes of 1 and 5 per week; by 1950 they were 3 and 4. Again for fats the pre-war figures showed extreme differences in consumption of 60 per cent., and for meat of 100 per cent., between the highest and the lowest income groups. By 1950 fats showed a negligible range and the differences in meat consumption were of the order of only 5 or 10 per cent. Only with fruit and vegetables were pre-war differences still observed. It is, of course, true that the levelling of consumption in the war and post-war years was due, in part at least, to measures specifically adopted to ensure an equitable distribution of supplies, based primarily on physiological needs—notably by rationing, subsidies, and welfare measures. Nevertheless the greater equalisation of incomes probably exerted the major effect. If so (and provided Welfare measures are maintained) this changed consumption pattern might well influence future levels of demand if foods were once again in free supply. It is here that the *third* factor will tend to operate—the future price of food. If consumption is to be maintained at both higher and physiologically more equitable levels, it is clear that the costs of production of our home supplies (and equally, of course, the costs of procurement of our overseas supplies) must be such as will furnish food at prices within the purchasing power of the mass of the population. This emphasises the importance of developing agricultural techniques which will be capable not only of increasing but of cheapening food production.

What of the availability of world supplies from overseas? The greatly increased demands caused by the immediate post-war needs of liberated Europe and the Far East have eased, but long-term tendencies may take their place. World population grows at a rate equivalent to 10 per cent. per decade. Moreover with increasing world consciousness of the importance of adequate nutrition, the more backward countries are naturally striving to improve their own living standards. Thus we may in the long term not merely find valuable pre-war sources of supply no longer open to us (oilseeds from India are a case in point), but we may well be faced with increasing competition for the food surpluses of the primary producing countries. What of the availability of currency? Here again the problem is not proving to be merely

short term, or merely one of dollar shortage. Owing to the war-time sacrifice of our interest-bearing assets overseas, combined with the expanding industrialisation of those nations to which we had previously looked for food (with its inevitable repercussions on the terms of trade) we are faced with an increasingly serious balance of payments problem, which will inevitably tend to limit our capacity to purchase our supplies from overseas.

With potentially increasing home demand (provided reasonable prices are maintained) and a potentially contracting availability of world supplies, the need for further home production is assured. Whether this will once again involve us in a sparser or a plainer diet cannot be forecast. It would, indeed, be idle to predict how far our present output could, if need be, be expanded, or to speculate upon its future form. But even within the present general pattern of land usage science, and the application of science, still offer ample scope for betterment. Crop yields on neighbouring farms still show too wide a gap ; plant and animal disease still take too great a toll, methods of breeding and of feeding still vary too greatly in efficiency and skill. It will, however, clearly remain to my successor in this chair, perhaps some ten or twenty years hence, to judge whether, within whatever pattern British agriculture has by then assumed, the industry has maintained its present march of progress, and whether in that march of progress science, the handmaiden of practice, has continued to fulfil her rightful and her beneficial role.[1]

[1] Delivered Sept. 5, 1952.

APPENDIX

TABLE I

IMPORTS OF FOOD AND ANIMAL FEEDS INTO THE UNITED KINGDOM

(Million tons)

	Pre-war	1940	1941	1942	1944	1947	1950
Wheat and other edible grains	6·9	7·4	6·6	4·0	3·7	5·4	4·7
Animal feeds	5·1	3·3	1·0	0·2	0·2	1·0	1·7
Sugar	2·2	1·5	1·7	0·8	1·2	1·9	2·1
Fruit and vegetables	2·5	1·4	0·3	0·3	0·4	1·7	1·9
Oilseeds and fats	2·5	2·3	2·2	2·1	2·1	1·7	2·1
Meat	1·5	1·5	1·6	1·7	1·8	1·4	1·3
Milk products and eggs	0·5	0·4	0·5	0·7	0·5	0·4	0·4
Other foods	1·3	1·4	1·2	1·0	1·2	1·3	1·5
Total	22·5	19·1	15·1	10·8	11·1	14·8	15·7

Sources : How Britain was fed in War-time, H.M.S.O., 1946, and current estimates for post-war years.

TABLE II

ESTIMATED ACREAGE AND PRODUCTION OF PRINCIPAL CROPS IN THE UNITED KINGDOM

	Pre-war	1940	1941	1942	1943	1944	1947	1948	1949	1950
ACREAGES (Million acres)										
Wheat	1·9	1·8	2·3	2·5	3·5	3·2	2·2	2·3	2·0	2·5
Potatoes	0·7	0·8	1·1	1·3	1·4	1·4	1·3	1·5	1·3	1·2
Sugar-beet	0·3	0·3	0·4	0·4	0·4	0·4	0·4	0·4	0·4	0·4
Barley	1·0	1·3	1·5	1·5	1·8	2·0	2·1	2·1	2·1	1·8
Oats	2·4	3·4	4·0	4·1	3·7	3·7	3·3	3·3	3·3	3·1
Fruit and vegetables	0·6	0·6	0·7	0·7	0·7	0·8	0·9	0·9	0·9	0·8
PRODUCTION (Million tons)										
Wheat	1·7	1·6	2·0	2·6	3·4	3·1	1·7	2·4	2·2	2·6
Potatoes	4·9	6·4	8·0	9·4	9·8	9·1	7·8	11·8	9·0	9·5
Sugar-beet	2·7	3·2	3·2	3·9	3·8	3·3	3·0	4·3	4·0	5·2
Barley	0·8	1·1	1·1	1·4	1·6	1·8	1·6	2·0	2·1	1·7
Oats	1·9	2·9	3·2	3·6	3·1	3·0	2·5	3·0	3·0	2·7
Fruit and vegetables	2·3	3·2	2·8	3·7	3·5	3·9	3·7	4·2	3·2	3·9

Sources : Annual Abstracts of Statistics 1938–1948 and 1938–1950.

TABLE III

ESTIMATED LIVESTOCK POPULATION AND PRODUCTION OF LIVESTOCK PRODUCTS IN THE UNITED KINGDOM

	Pre-war	1940	1941	1942	1943	1944	1947	1948	1949	1950
POPULATION (Millions)										
Cattle	8·8	9·1	8·9	9·1	9·3	9·5	9·6	9·8	10·2	10·6
Sheep	26·8	26·3	22·3	21·5	20·4	20·1	16·7	18·2	19·5	20·4
Pigs	4·4	4·1	2·6	2·1	1·8	1·9	1·6	2·2	2·8	3·0
Poultry	74·2	71·2	62·0	57·8	50·7	55·1	70·0	85·4	95·5	96·1
PRODUCTION (Thousand tons except milk)										
Milk (million gallons)	1·2	1·2	1·2	1·3	1·3	1·4	1·5	1·6	1·7	1·9
Beef	605	582	551	481	478	519	510	499	528	636
Mutton	211	227	176	171	159	142	117	123	140	147
Pigmeat	393	362	248	150	150	169	113	136	232	290
Eggs	389	430	417	314	266	264	313	358	423	516

Sources : Annual Abstracts of Statistics 1938–1948 and 1938–1950.

TABLE IV

FOOD CONSUMPTION LEVELS IN THE UNITED KINGDOM

(lb. per head per year)

	Pre-war	1941	1944	1947	1950
Dairy products	38	41	49	49	54
Meat *	129	99	110	94	112
Fish, poultry and game *	35	22	25	38	30
Eggs and egg products	29	19	23	21	32
Oils and fats *	51	44	43	38	51
Sugar	110	78	76	85	80
Flour	195	237	234	224	202
Potatoes	176	188	275	284	244
Fruit and vegetables	249	169	218	246	228

* Product weight.

Sources : Food Consumption Levels in the United Kingdom, Cmd. 7842.
Ministry of Food Bulletins No. 611 and 655.

TABLE V

ESTIMATED NUTRIENT EQUIVALENT OF UNITED KINGDOM FOOD SUPPLIES

(per head per day)

		Pre-war	1941	1944	1947	1950
Calories		3,000	2,820	3,010	2,880	2,990
Total protein	g.	80·4	82·4	87·1	90·7	88·8
Animal protein	g.	43·2	38·3	41·4	43·0	46·0
Calcium	mg.	695	698	1,039	1,190	1,207
Iron	mg.	12·6	12·9	16·2	16·6	15·8
Vitamin A	i.u.	4,042	3,604	3,799	3,840	4,345
Vitamin B_1 *	mg.	1·2	1·5	2·0	1·8	1·8
Vitamin C	mg.	96	81	106	112	100

* Other constituents of the vitamin B complex (i.e. riboflavin and nicotinic acid) increased equivalently.

Sources : Food Consumption Levels in United Kingdom, Cmd. 7842.
Ministry of Food Bulletins No. 611 and 655.

TABLE VI

PERCENTAGE CHANGES IN THE CONTRIBUTION OF BRITISH AGRICULTURE TO UNITED KINGDOM SUPPLIES

(Pre-war = 100)

	Pre-war	1940	1942	1943	1944	1947	1948	1949	1950
Calories *	100	103	116	133	130	130	119	126	133
Animal protein	100	94	84	82	81	95	106	106	108
Net output †	100	108	105	120	125	117	123	135	140

* Based on the pre-war contribution of 31 per cent.
† Based on produce from agricultural holdings at 1945–46 prices.

Sources : *How Britain was fed in War-time*, H.M.S.O., 1946.
Agricultural Statistics, United Kingdom, Part II, 1939–1946 and current Estimates for post-war years.

Printed in Great Britain
SPOTTISWOODE, BALLANTYNE & CO. LTD.
London & Colchester